Investigation of Hallmark Epigenetic Changes in a Cancer Stem Cell Model

Laurence Wild MA (Cantab)

University College London

2010

Submitted for the degree of Doctor of Philosophy

U.P : up

Declaration

I, Laurence Wild, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Laurence Wild

November 2009

Abstract

Epigenetic control of gene expression is vital for normal development and differentiation of cells, and is also important in the development of disease. In particular, there is a strong association between hallmark epigenetic changes and cancer – namely, genome wide hypomethylation, gene specific hypermethylation and characteristic histone modification. Almost all studies of cancer epigenetics to date have been conducted in malignant tissues or already transformed cell lines, and therefore do not take into account epigenetic changes occurring *during* the process of transformation. Our lab has developed a line of primary mesenchymal stem cells (MSC; thought to be the origin of various types of sarcoma) in which five oncogenic steps towards a fully transformed state are sequentially introduced including: human telomerase, necessary to extend the life span of MSC in culture, genes to inactivate the p53 and pRb tumour suppressor genes and genes to activate the oncogenes c-Myc and Ras. I hypothesized that hallmark epigenetic changes take place in this step-wise model of transformation, and aimed to investigate genome wide hypomethylation and the activity of the polycomb repressive 2 (PRC2) complex in this system. Utilizing the PCR based technique MethyLight, I show that transformed MSC are hypomethylated compared to parental MSC, with this decrease in methylation occurring on the introduction of oncogenic H-Ras in the final step. I also show that this hypomethylation is a gradual event following H-Ras expression, and transformation can take place in the absence of hypomethylation. I demonstrate that the three core components of the PRC2 complex are up-regulated during step-wise transformation and that PRC2 target genes are down-regulated. Finally, I show that MSC are able to be transformed when the PRC2 components EZH2 and SUZ12 are knocked down before the final oncogenic hit. These studies show that hallmark epigenetic changes occur during step-wise transformation and suggest that tumour-associated epigenetic changes occur following genetic aberrations. This model is valuable and relevant to further explore the mechanisms behind epigenetic alterations in cancer.

Acknowledgements

First and foremost I would like to thank my PhD supervisors James Flanagan and Chris Boshoff for their support and guidance over the last four years.

I am extremely thankful to the members of the CR-UK Viral Oncology group for help and advice in the lab. In particular, my thanks go to Juanma Funes for assistance with the cell lines and Leonid Nikiteno for guidance with western blotting. Thanks also to the MRC for funding my studentship.

Finally, I am grateful beyond words to Abby and my parents for their ever-present support and encouragement.

Publications

Publications arising from the work described in this thesis at the time of submission:

1. **Wild, L.**, Funes, J.M., Boshoff, C., and Flanagan, J.M. "*In vitro* transformation of mesenchymal stem cells induces passive genomic hypomethylation." *In press, Carcinogenesis*
2. **Wild, L.** and Flanagan, J.M. "Genome-wide hypomethylation in cancer: how and when does it occur?" *In press, B.B.A. Reviews on Cancer*
3. Flanagan J.M. Funes J.M., **Wild L.,** Carey N., Henderson S., and Boshoff C. Genomics screen in transformed stem cells reveals RNASEH2A, PPAP2C and ADARB1 as putative anti-cancer drug targets. *Mol. Cancer Ther*. 2009 Jan;8(1):249-60.
4. Flanagan J.M. and **Wild L**. An epigenetic role for noncoding RNAs and intragenic DNA methylation. *Genome Biol.* 2007 Jun 27;8(6):307.

Abbreviations

4-OHT	4-hydroxytamoxifen
AML	Acute myeloid leukaemia
AZA	5-aza-2'-deoxycytidine
bFGF	Basic fibroblast growth factor
bp	Base pair
BSA	Bovine serum albumin
°C	Degrees celcius
CpG	Cytidine-guanosine dinucleotide
ChIP	Chromatin immunoprecipitation
ChIP-chip	ChIP followed by microarray hybridization
ChIP-seq	ChIP followed by next generation sequencing
CML	Chronic myeloid leukaemia
Ct	Threshold cycle
ddH_2O	Double distilled water
DMEM	Dulbecco's modified Eagle medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
dNTP	2'-deoxyribonucleoside 5'-triphophate
DMR	Differentially methylated region
EDTA	Ethylene diamino tetraacetic acid
ERK	Extracellular signal-regulated kinase
ER^{TM}	Mutated oestrogen receptor
EV	Empty vector control
FACS	Fluorescence assisted cell sorting
FBS	Foetal bovine serum
FDR	False discovery rate

FITC	Fluorescein isothiocyanate
FL1 / FL2	Fluorescence channels for FACS
Gag-pol	Group specific antigen and polymerase
GEM	Gene expression microarray
GFP	Green fluorescent protein
GSEA	Gene set enrichment analysis
H3	Histone H3
H1K26	Histone H1 lysine residue 26
H3K9	Histone H3 lysine residue 9
H3K27me3	Trimethylation of histone H3 lysine residue 27
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HF	Human fibroblast
hr	Hour
HPLC	High performance liquid chromatography
HPV	Human papilomavirus
HRP	Horseradish peroxidase
HSP	Heat shock protein
hTERT	Catalytic subunit of human telomerase
ICF	Immunodeficiency, centromeric instability and facial abnormality syndrome
IgG	Immunoglobulin G
IP	Immunoprecipitation
Kb	Kilobase pairs
LB	Luria-Bertani broth
LINE1	Long Interspersed Nuclear Element 1
MAPK	Mitogen-activated protein kinase
miRNA	microRNA
MBD	Methyl binding domain

MeDIP	Methylated DNA immunoprecipitation
min	Minutes
MPNST	Malignant peripheral nerve sheath tumour
mRNA	Messenger RNA
MSC	Mesencymal stem cell
Min	Minute
MSP	Methylation specific PCR
MTS	Cell viability assay
ncRNA	Non-coding RNA
NF	Neurofibroma
NS	Non-silencing shRNA control
PBS	Phosphate buffered saline
p.i.	Post infection
pERK	Phosphorylated ERK
PcG	Polycomb group
PCR	Polymerase chain reaction
PI	Propidium iodide
PI3K	Phosphatidylinositol 3-kinase
PP2A	Protein phosphatase 2A
PRC	Polycomb repressive complex
qPCR	Quantitative polymerase chain reaction
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
Ral-GEF	Ral-guanine nucleotide exchange factor
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
RNAi	RNA interference
SAM	S-adenosyl methionine
Sat-α	Satellite alpha pericentromeric repeat

Sat2	Satellite-2 pericentromeric repeat
sec	Second
SEM	Standard error of the mean
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
SV40	Simian virus 40
TAE	Tris-acetate ethylene diamino tetraacetic acid
TEMED	Tetramethylethylenediamine
Tris	Tris(hydroxymethyl)aminomethane
TSS	Transcription start site
U	Units
UTR	Untranslated region
v/v	Volume / volume
VSV-G	Vesicular stomatitis virus glycoprotein
w/v	Weight / volume
WB	Western blot
wrt	With respect to
YFP	Yellow fluorescent protein

Table of Contents

List of figures

Chapter 1

Chapter 3

Chapter 4

Chapter 5

List of tables

Chapter 2

Chapter 5

Appendix

Chapter 1: Introduction

1.1 Epigenetic Regulation of Gene Expression

'Epigenetics' describes the study of stable, reversible alterations affecting gene expression that are not stored within the primary DNA sequence. Since the identification of epigenetic regulation, there has been rapid progress in elucidating the nature and role of the mechanisms involved in this type of gene expression control (Qiu, 2006). Such research has shown that epigenetic factors play a vital role not only in development and differentiation, but also in pathogenesis. In particular, a strong link between epigenetic changes and the development of cancer has begun to emerge. In order to further explore this role of epigenetics in cancer development, we must first consider the nature of epigenetic control acting across the genome.

1.1.1 DNA Methylation

Nucleic acids can be modified by the covalent addition of a methyl group, through interaction with an alkylating agent or by an enzymatic reaction. DNA methylation in mammalian genomes occurs predominantly on cytosine residues 5' of guanine in CpG dinucleotides (Jaenisch and Bird, 2003). Approximately 70% of CpG dinucleotides in the human genome are methylated, though the distribution of CpG dinucleotides and frequency of methylation are not random (Cooper and Krawczak, 1989). Although the CpG sequence is underrepresented

in the genome as a result of spontaneous deamination of methylated cytosine to thymine, CpG dinucleotides are highly abundant in stretches approximately 1kb in length known as CpG islands – regions often associated with the promoters and first exons of genes (Bird et al., 1985). Despite constituting only 1% of the genome and containing 15% of CpG dinucleotides, 50% of unmethylated CpG sites can be found in CpG islands. These regions tend to have open chromatin structures lacking the linker histone H1 and are enriched for acetylated histone H3 and H4 (Tazi and Bird, 1990).

DNA methylation is established and maintained by three DNA methyltransferases (DNMTs) – DNMT1, DNMT3a and DNMT3b. DNMT1, the first of these enzymes to be identified, is ubiquitously expressed, localises to replication foci and exhibits a strong preference for hemi-methylated DNA (Leonhardt et al., 1992; Pradhan et al., 1999; Robertson et al., 1999). These observations indicate that the enzyme is involved in the copying of DNA methylation to the daughter strand following DNA replication, leading to the common classification of DNMT1 as a 'maintenance' methyltransferase. DNMT1 is vital for normal embryonic development, imprinting and X-inactivation, with knockout mice showing global demethylation and embryonic lethality (Lei et al., 1996; Li et al., 1992).

In contrast, the DNMT3a and DNMT3b enzymes are highly conserved from zebrafish to humans and show an affinity for both hemi- and fully unmethylated DNA, suggesting a potential role in the *de novo* establishment of DNA methylation (Okano et al., 1998). In line with this, these enzymes are required for *de novo* genomic DNA methylation following embryo implantation and for the silencing of newly integrated retroviral sequences in mouse ES cells

(Okano et al., 1999). *Dnmt3a* knockout mice are born live but die within 4 weeks, whereas *dnmt3b* knockout is embryonic lethal around the same point as *dnmt1* knockout, embryonic day (E) 9.5 (Li et al., 1992; Okano et al., 1999).

However, the classification of these enzymes as 'maintenance' or 'de-novo' methyltransferases may not be so clear cut. Overexpression of DNMT1 in cancer cells has been shown to induce de-novo methylation of certain genes, although this effect may be the result of DNMT1 replicating methylation marks established by other methyltransferases (Vertino et al., 1996). A further study has demonstrated that cancer cells lacking *DNMT1* maintain 80% of their normal methylation patterning, suggesting that DNMT3A and DNMT3B may also act as maintenance methyltransferases (Rhee et al., 2000).

1.1.2 Histone Modifications and Chromatin Structure

Chromatin structure also plays an important role in the regulation of transcription by controlling access of transcriptional machinery to promoter regions. Chromatin is comprised of DNA wrapped around a series of nucleosomes - octomers made up of pairs of the histones H2A, H2B, H3, and H4 (Luger et al., 1997; Luger and Richmond, 1998). Histones themselves are made up of a globular protein domain and an N-terminal tail, which can be the target of various modifications including methylation, acetylation, phosphorylation, ubiquitination and sumoylation. Such modifications alter histone structure, causing changes to chromatin conformation and affecting accessibility of the region to transcriptional machinery (Jenuwein and Allis, 2001; Turner, 2002).

Of the numerous modifications that can occur to histone tails, acetylation and methylation are the most common. Histone acetylation is mediated by histone deacetylases (HDACs) and histone acetyltransferases (HATs). HATs acetylate the ε-NH2 group on lysine residues within histone tails, causing an open chromatin conformation and transcriptional activation (Marmorstein and Roth, 2001; Sterner and Berger, 2000). Conversely, HDACs can cause chromatin condensation and transcriptional repression by removing acetyl groups on histone tails (Thiagalingam et al., 2003).

Methylation of histone tails by histone methyltransferases (HMTs) also plays an important role in regulating chromatin structure. These methylation events can take place at a number of different residues (including lysine and arginine), with each change having a different effect. For example, methylation of lysine 9 and lysine 27 in the H3 tail is associated with transcriptionally repressed chromatin, whereas H3 lysine 4 methylation is a characteristic of transcriptionally active chromatin (Zhang and Reinberg, 2001).

Finally, chromatin structure can also be altered through the action of ATP-dependant chromatin remodelling complexes such as SWI/SNF2/Brm, ISWI, and the Mi-2/NuRD complex (Havas et al., 2000). Polycomb-group and trithorax group genes also play roles in chromatin remodelling, acting as epigenetic silencers and activators for gene transcription and will be discussed in greater depth later.

1.1.3 MicroRNAs

MicroRNAs (miRNAs) are small non-coding RNAs that function as negative gene regulators through one of two mechanisms. First, miRNAs may function through the RNA-mediated interference (RNAi) pathway, leading to cleavage of complementary mRNA molecules. Alternatively, interactions with the 3' untranslated region (UTR) of target mRNA can lead to post transcriptional repression, decreasing protein levels without affecting the amount of mRNA (Esquela-Kerscher and Slack, 2006).

1.2 The Role of DNA Methylation

1.2.1 Repression of Repetitive Element Transcription

In considering the evolutionary role of DNA methylation, it is worth noting the genomic positioning of methylated CpG dinucleotides. Although much attention has focused on DNA methylation in CpG islands, the majority of methylated CpG dinucleotides are in fact found within repetitive elements, in particular parasitic DNA elements such as transposons, which in total comprise approximately 45% of the human genome (Lander et al., 2001). Transposable elements pose a significant threat to the integrity of the genome in a number of ways. Insertion of such elements during transposition can disrupt coding sequences of genes or result in aberrant initiation sites (Kazazian and Moran, 1998; Montagna et al., 1999). It is believed that DNA methylation is an essential mechanism in silencing the transcription of these elements to prevent their movement and expansion throughout the genome (Yoder et al., 1997). In line with this, it has been shown that methylation of retrotransposon promoters results in repression of their transcription (Kochanek et al., 1995).

In line with this, a number of studies have noted increased transcription of repetitive elements following genomic hypomethylation. *Dnmt1* knockout mice, which retain only 30% of normal genomic methylation, show significantly increased expression of the intracisternal A particle (IAP) repetitive element (Walsh et al., 1998). It is believed that increased transposition of this element may underlie the genomic instability and frequent gene rearrangements observed in these mice (Chen et al., 1998). Kangaroo hybrid models, which display

genome wide hypomethylation, also show expansion of retroviral like elements (O'Neill et al., 1998). Similarly, global loss of DNA methylation is a common event in a number of cancers, and can result in aberrant transcription of endogenous transposable elements, such as elevated LINE1 transcription in CML (Roman-Gomez et al., 2008).

1.2.2 Transcriptional Repression

DNA methylation is also traditionally recognized as a repressor of gene expression, and can function in this regard through two main mechanisms – by interfering with the assembly of the transcription machinery or by causing alterations to chromatin structure. The transcriptional repressor methyl-CpG binding protein 2 (MECP2), which can bind to methylated DNA through its methyl-CpG binding domain (MBD) whilst recruiting co-repressors through the transcriptional repressor domain (TRD), plays an important role in both of these mechanisms. *In vitro* experiments have shown that MECP2 represses transcription from methylated templates through interactions between the TRD and TFIIB, a key component of the transcriptional machinery (Kaludov and Wolffe, 2000). It is believed that MECP2 functions in a similar manner *in vivo*, as it has been shown to cluster at sites of DNA methylation (Nan et al., 1996).

Methylated DNA binding proteins, including MECP2, MBD1, MBD2 and MBD3, also play a role in linking DNA methylation to the formation of a repressive chromatin structures. For example, MECP2 associates with repressor complexes including histone deacetylases - enzymes able to induce a stable

repressive chromatin state - with this gene repression partially sensitive to the histone deacetylase inhibitor trichostatin A (TSA) (Jones et al., 1998; Nan et al., 1998). Furthermore, MBD3 has been shown to associate with the Mi-2 chromatin remodelling complex in *Xenopus* eggs, and MBD2 can be co-purified with histone deacetylases (HDAC) 1 and 2 in HeLa cell nuclear extracts (Ng et al., 1999; Wade et al., 1999).

As well as these interactions mediated by methyl-CpG binding proteins, direct associations have been observed between enzymes involved in establishing DNA methylation and those involved with chromatin modification. For example, the minimal interaction domain of DNMT1 has been shown to mediate associations between this enzyme and histone deacetylases, leading to gene repression that can be partly relieved by TSA treatment (Fuks et al., 2000; Robertson et al., 2000; Rountree et al., 2000).

1.2.3 Maintenance of Genomic Integrity

As well as its role in the regulation of gene and repetitive element expression, DNA methylation plays a vital role in the maintenance of genomic integrity through repression of recombination events between homologous repetitive sequences across the genome. Studies in the fungus *Ascobolus immersus* have demonstrated a reduction in recombination events when known recombination hotspots are methylated (Maloisel and Rossignol, 1998). Similar results have been obtained in studies of V(D)J recombination in mammalian systems (Hsieh and Lieber, 1992). Other lines of evidence in support of this hypothesis come

from studies of genome wide loss of methylation in disease and following drug treatment or *in vitro* genetic manipulation of cells. Mouse embryonic stem cells deficient in *dnmt1* show a significant increase in abnormal gene rearrangements (Chen et al., 1998). Similarly, ICF (immunodeficiency, centromeric instability, facial anomalies) syndrome patients display characteristic hypomethylation of satellite repeats alongside chromosomal recombination events that can be recapitulated in cell lines following treatment with the demethylating agent 5-aza-2'-deoxycytidine (Jeanpierre et al., 1993; Ji et al., 1997; Tuck-Muller et al., 2000). Finally, genome-wide hypomethylation is a common event in cancer and could lead to an increase in abnormal karyotypic changes in transformed cells (Ehrlich, 2002).

1.3 Epigenetic Deregulation In Cancer

The six hallmarks of cancer necessary for continued cell growth are: self sufficiency in growth signals, insensitivity to antigrowth signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis and metastasis (Hanahan and Weinberg, 2000). Dominant gain of function mutations, or other increases in the transcription of genes involved in these pathways, can contribute towards cellular transformation. Furthermore, cells must avoid inbuilt anticancer defence mechanisms through loss of tumour suppressor function in order for transformation to take place. Although many of these changes occur at the DNA level through mutation, there is mounting evidence that epigenetic changes in transcription are also strongly associated with transformation of cells. Such changes could affect DNA methylation, chromatin structure, or both, leading to increased expression of oncogenes, or reduced expression of tumour suppressors.

Numerous studies have documented epigenetic changes in cancer cells including genome-wide hypomethylation, gene specific hypermethylation and characteristic histone modifications (Esteller, 2005; Feinberg and Tycko, 2004; Fraga et al., 2005). Genome-wide hypomethylation has been widely documented in cancer and is associated with loss of DNA methylation at repetitive and structural elements across the genome, potentially leading to genomic instability (Ehrlich, 2002). Hypermethylation, however, takes place in promoter regions of tumour-suppressor genes such as $p16^{INK4a}$, *BRCA1* and the mismatch repair gene *hMLH*, resulting in gene silencing (Esteller, 2005; Jones and Baylin, 2002). As well as tumour suppressors, promoter regions for a broad range of other genes have been found to be hypermethylated in cancer cells. These include genes

involved in signal transduction (*RASSF1* and *APC*), carcinogen metabolism (*GSTP1*), apoptosis (*DAPK* and *CASP8*), hormone response (*EG, PGR, AR*), angiogenesis (*THBS1*) and metastasis (*TIMP3*, *CDH1*) (Kopelovich et al., 2003). Genes involved in the establishment and maintenance of epigenetic changes are also deregulated in cancer. Such genes include DNA methyltransferases, histone acetyltransferases, histone deacetylases, methyl-CpG binding proteins, chromatin remodelling factors and SWI/SNF family proteins (reviewed in Esteller, 2006).

In addition to methylation changes, histone modifications also play a key role in the regulation of gene expression. Certain histone modifications, such as histone 3 lysine 9 (H3K9) and 27 (H3K27) methylation have been shown to be involved in aberrant gene silencing in cancer (Feinberg et al., 2006). Additionally, it has been observed that tumours display 'hallmark' histone modifications, including loss of acetylation at K16 and trimethylation at K20 of histone H4 (Esteller, 2006; Feinberg and Tycko, 2004; Fraga et al., 2005).

Finally, ATP dependant chromatin remodelling complexes such as the SWI/SNF complex and the antagonistic polycomb family of proteins can also affect nucleosome structure. The interplay between these proteins, DNA methylation changes and histone modifications all act to affect gene expression thorough changes in higher order chromatin structure. Alterations to the activity of such proteins could, therefore also play a large role in the epigenetic changes associated with cancer (Esteller, 2006).

MicroRNAs have been shown to be involved in the regulation of cell growth, differentiation and apoptosis – pathways commonly deregulated in transformed cells. In line with this, miRNA misexpression and mutation are associated with various cancers. Furthermore, genes involved in the biogenesis

and processing of miRNAs can themselves be considered as epigenetic regulators of gene transcription, and as such may play a role in the transformation process (Esquela-Kerscher and Slack, 2006).

Despite significant progress in the field of cancer epigenetics, the timing and causes of these changes remain subjects of investigation. Several studies have suggested that epigenetic changes occur early in the transformation process (Fearon and Vogelstein, 1990; Feinberg et al., 2006; Suter et al., 2004). However, most epigenetic studies of cancer are performed on malignant tissues, meaning that any observed epigenetic changes could be a consequence of the disease rather than a cause, or represent non-specific changes unassociated with the malignancy.

1.4 Genome Wide DNA Hypomethylation in Cancer

Many cancers display hallmark epigenetic changes including characteristic histone modifications, gene specific hypermethylation and global hypomethylation (Feinberg and Tycko, 2004; Fraga et al., 2005). Of these, the first to be described was genome wide hypomethylation, a decrease in the overall genomic methylcytosine content from approximately 4% in normal tissues to 2-3% in cancers. This change was first observed in a 1983 study comparing lung and colon carcinomas to adjacent normal tissue, demonstrating that overall genomic methylcytosine levels were lower in cancer tissues (Feinberg and Vogelstein, 1983). Since this pioneering study, the observation has been reproducibly repeated in comparisons of a wide range of cancers and normal tissues using a variety of different techniques (Table A1; Ehrlich et al., 2002; Feinberg et al., 1988; Gama-Sosa et al., 1983b; Narayan et al., 1998; Qu et al., 1999a; Weisenberger et al., 2005).

1.4.1 Genomic Locations of Hypomethylation

The precise genomic locations at which methylation is lost also remains a subject of investigation. An early study of methylation in randomly fragmented DNA from human gastric adenocarcinoma demonstrated that hypomethylation was present in highly repetitive, moderately repetitive and single copy DNA sequences across the genome (Gama-Sosa et al., 1983b). However, the magnitude of the decrease in methyation observed in cancers is most commonly

accounted for by loss of methylation at repetitive sequences, which constitute ~45% of the genome as a whole – comprising interspersed repeats derived from non-autonomous or autonomous transposable elements and tandem repeats (Ehrlich, 2002; Lander et al., 2001; Prak and Kazazian, 2000).

One of the most widely studied repetitive elements is the Long Interspersed Nuclear Element 1 (LINE-1), a 6kb interspersed DNA repeat which makes up around 15% of the human genome. Instances of LINE-1 hypomethylation have been reported in chronic lymphocytic leukaemia, hepatocellular carcinoma, prostate carcinoma and urinary bladder carcinoma amongst others (Dante et al., 1992; Jurgens et al., 1996; Santourlidis et al., 1999; Takai et al., 2000).

Another highly repetitive interspersed DNA sequence is the ~300bp Alu repeat, with an average copy number of over 10^6, constituting approximately 10% of the human genome. However, changes in methylation of Alu repeats have not been widely investigated, with the exception of one study in germline tumours (Rubin et al., 1994). Moderately repeated DNA sequences such as human endogenous retrovirus-K (HERV-K; 30-50 full length sequences and 10000 LTRs in the human genome) have also been observed to be hypomethylated in urinary bladder cancer (Florl et al., 1999).

Satellite DNAs, highly repetitive tandem sequences commonly found in the heterochromatic regions around the centromere, also constitute a common site of hypomethylation in cancer. Satellite alpha (Sat α) and Satellite 2 (Sat 2), located primarily on chromosomes 1, 10 and 16, have been extensively studied and shown to be hypomethylated in breast adenocarcinoma, ovarian epithelial

carcinoma and Wilms tumours, among others (Narayan et al., 1998; Qu et al., 1999b).

Although global hypomethylation is often accounted for by the loss of methylation across these repetitive elements, single copy sequences in the genome including imprinted genes also exhibit hypomethylation in cancer. Examples include the gamma-globin gene in breast and colon adenocarcinomas, *MN/CA9* in renal cell carcinomas and *c-MYC* in colorectal cancers and hepatocellular carcinomas (Cho et al., 2001; Ribieras et al., 1994; Sharrard et al., 1992; Shen et al., 1998).

1.4.2 Appropriate Normal Controls for Studies of Hypomethylation

The variation in global methylation levels between different cells and tissues must be taken into account when choosing appropriate cases and controls for studies of global hypomethylation in cancer (Gama-Sosa et al., 1983a; Gama-Sosa et al., 1983b). The variable nature of DNA methylation between tissues means that appropriate controls must always be used. At the extremes, a comparison of global methylcytosine content in normal somatic tissues demonstrated a difference of 17% between cerebellum (4.03 +/- 0.14%) and heart (3.43 +/- 0.12%) (Ehrlich et al., 1982). Furthermore, these tissues represent a mixed population of cells that may not be the same as the tumour population. In order to account for this variation, many studies assess tumour tissue methylation in comparison to matched adjacent normal tissue or a panel of unmatched normal tissues.

1.4.3 Timing of Genome-wide Hypomethylation

The timing of global DNA hypomethylation in cancer has also been a subject of investigation. Since global DNA hypomethylation has been observed in most cancers assayed, as well as in benign tumours, it is believed to constitute an early event in transformation (Fearon and Vogelstein, 1990; Feinberg et al., 2006). This hypothesis is supported by a study of LINE1 promoter methylation (a marker of global methylation levels) in colorectal cancer patients, which observed global hypomethylation in tumours and apparently normal adjacent tissue when compared to control tissue from healthy individuals (Suter et al., 2004). This hypomethylation was not present in peripheral blood from the same patients, suggesting that the loss of methylation was an early event that occurred in normal tissues before the acquisition of further oncogenic changes.

The relative timings of global hypomethylation and gene specific hypermethylation in cancer have also been utilized to address whether hypomethylation occurs at earlier stages of the transformation process. Although many studies have failed to elucidate a link between the two events, an investigation of LINE1 methylation levels in prostate carcinomas of different stages found that samples exhibiting hypomethylation formed a subclass of those with gene specific hypermethylation. The lack of tumours exhibiting hypomethylation in the absence of gene specific hypermethylation suggests that global loss of methylation is a secondary event in the progression of this cancer (Florl et al., 2004). Similarly, studies of genome-wide hypomethylation and gene specific hypermethylation in prostate cancer progression have suggested that hypomethylation only occurs in late stage metastatic cancer (Yegnasubramanian

et al., 2008; Yegnasubramanian et al., 2004). Further investigations are required to establish the relative timings of these epigenetic changes during transformation (reviewed in Ehrlich, 2002; Wilson et al., 2007).

1.4.4 Consequences of Genome Wide Hypomethylation

As discussed previously, a large proportion of DNA methylation in the genome is associated with repetitive elements, including many types of transposable elements as well as tandem repeats such as satellite sequences. The strong sequence similarity between repetitive elements can increase the probability of recombination events taking place between these loci, with this recombination normally suppressed in eukaryotes through DNA methylation and subsequent formation of heterochromatin, restricting access of these sequences to recombination machinery (Garrick et al., 1998; Nguyen et al., 2001). It is believed that loss of methylation at such sequences, as is often observed in cancer, could reverse this restriction and contribute to karyotypic instability and tumour progression. In order to investigate this possibility, parallels have been drawn between DNA hypomethylation in cancer and ICF (immunodeficiency, centromeric instability, facial anomalies) syndrome. Hypomethylation of pericentromeric Sat2 repeats on chromosomes 1 and 16 is commonly observed in cancer and is also a characteristic of ICF. In karyotypic studies of ICF lymphocytes and lymphoblastoid cell lines, these regions of chromosomes 1 and 16 are specifically targeted for chromosomal rearrangements (Jeanpierre et al., 1993; Tuck-Muller et al., 2000). Additionally, Sat2 rich regions are often the site

of chromatin decondensation and the break point in chromosomal deletions or in the formation of multi-radial chromosomes in these cells (Ehrlich et al., 2001; Tuck-Muller et al., 2000). Finally, global DNA hypomethylation induced by 5-azacytidine treatment of normal lymphoid cultures results in karyotypic changes identical to those observed in ICF (Hernandez et al., 1997; Ji et al., 1997).

Similar pericentromeric rearrangements involving chromosomes 1 and 16 have been described in both hepatocellular carcinoma, breast cancer and Wilms tumour, and have been shown to correlate with hypomethylation of Sat2 sequences on these chromosomes (reviewed in Ehrlich, 2002; Tsuda et al., 2002). As well as karyotypic instability resulting from hypomethylation of pericentromeric sequences, global DNA hypomethylation alongside an increase in abnormal chromosomal rearrangements has been observed in *dnmt1* knockout mouse embryos as well as in a mammalian hybrid model (Biniszkiewicz et al., 2002; Chen et al., 1998; O'Neill et al., 1998). Similar observations have also been made in human cancers, including increased chromosomal instability in conjunction with genome wide hypomethylation in hepatocellular and prostate carcinomas as well as colorectal cancer cell lines and tumours (Karpf and Matsui, 2005; Lengauer et al., 1997; Rodriguez et al., 2006; Schulz et al., 2002; Wong et al., 2001). Together, these results strongly suggest that global hypomethylation, particularly of Sat2 elements, can lead to karyotypic abnormalities in cancer through increased frequency of recombination events.

As well as promoting karyotypic instability, the strong link between DNA methylation and transcriptional silencing suggests that hypomethylation in cancer may lead to elevated transcription of genes which favour oncogenesis. Cancer associated increases in gene expression due to promoter hypomethylation

have been reported for genes including *PS2, HOX11 and MN/CA9* (Cho et al., 2001; Martin et al., 1997; Watt et al., 2000) and other proto-oncogenes (Tao et al., 2000). Similarly, global hypomethylation may contribute to loss of silencing at imprinted loci, leading to aberrant expression of paternal transcripts (Sullivan et al., 1999; Takai et al., 2001). Finally, studies of global hypomethylation in mammalian model systems have also demonstrated increased expression of transposons as a result of demethylation of these elements. For example, elevated expression of the instracisternal-A particle (IAP) retrotransposon has been reported in *dnmt1* deficient mouse embryos (Walsh et al., 1998). Similarly, demethylation of the LINE1 promoter in CML has been shown to increase transcription of this element (Roman-Gomez et al., 2008).

Hypomethylation across the genome may also alter gene expression patterns through *trans* effects. In particular, demethylation may disrupt the normal heterochromatin-euchromatin balance in cells, as suggested in ICF lymphoid cells (Ehrlich 2001). Alternatively, aberrant recruitment of proteins involved in transcriptional control to demethylated regions of the genome may disrupt the action of these proteins in normal transcription (reviewed in Ehrlich, 2002).

1.4.5 Potential Causes of Genome Wide Hypomethylation

1.4.5.1 Passive DNA Demethylation

The importance of DNA methyltransferases (DNMTs) in establishing and maintaining DNA methylation patterns suggests that disruption of the activity of these genes could result in passive global hypomethylation through loss of methylation on consecutive rounds of DNA replication. For example, *dnmt1* knockout mice display global hypomethylation and increased frequency of certain cancers including T-cell lymphomas, hepatocellular adenomas and carcinomas (Gaudet et al., 2003; Yamada et al., 2005). However, it is unclear whether reduced DNMT1 expression leads to genome-wide hypomethylation in cancer. Studies of DNMT1 levels in a variety of tumour types show that expression correlates with cell proliferation, but there is no significant relationship between DNMT1 expression and incidence of hypomethylation in cancer (reviewed in Eads et al., 1999; Wilson et al., 2007).

Mutations in DNMT3B have been linked to hypomethylation of the Sat2 tandem repeat in ICF patients (Hansen et al., 1999; Xu et al., 1999). While similar patterns of hypomethylation at this element are observed in cancer, no study has yet linked this to reduced DNMT3B expression or function. In fact, increased DNMT3B levels in relation to other DNMTs have been reported in several cancers (Girault et al., 2003; Kanai et al., 2001).

The interplay between histone modifications and DNA methylation also suggests that changes to these marks in transformed cells may affect the global DNA methylation profile and could potentially play a role in the induction of

hypomethylation. In line with this, an investigation of histone modifications in a variety of cancer cell lines and tumours revealed loss of acetylation of H4K16 and increased trimethylation of H4K20 in cancers compared to normal controls. Furthermore, these changes were found to occur at repetitive elements – such as the Sat2 pericentromeric repeat as well as subtelomeric D4Z4 and NBL2 repeats – and correlated with hypomethylation of these regions. However, the study could not establish whether changes in these histone modifications were a cause or consequence of the DNA hypomethylation at these sites (Fraga et al., 2005).

1.4.5.2 Active DNA Demethylation

Although the mechanisms described above could result in a 'passive' loss of DNA methylation through disruption of maintenance or *de-novo* methylation, DNA hypomethylation could alternatively be induced in an 'active' manner through direct removal of the methyl groups. Whilst the enzymes involved in establishing DNA methylation have been well characterised, the mechanisms by which methyl groups are removed from cytosine residues are less well explored.

In particular, genome wide changes in methylation during gametogenesis and early embryogenesis provide the strongest evidence in favour of an active demethylation mechanism. In mammals, there are at least two periods during which global DNA methylation patterns are drastically altered – during zygote development and in the preimplantation embryo - in order to impart somatic cells with the epigenetic reprogramming required for the broad developmental

potential and appropriate imprinting of gametes and the early embryo (reviewed in Morgan et al., 2005; Reik et al., 2001; Santos et al., 2002).

In mice, the genomes of both male and female primordial germ cells undergo global DNA demethylation around embryonic day 13 or 14. Following this period, the cells enter mitotic or meiotic arrest, after which methylation patterns are re-established (reviewed in Reik et al., 2001). It is not clear, however, whether this process is an active or passive demethylation.

Clearer evidence for active demethylation occurs in the early embryo following fertilization. During this period, the male pronucleus undergoes drastic modification, including replacement of protamines with histones and global DNA demethylation. Crucially, this demethylation occurs before DNA replication is complete, indicating that the loss of methylation is an active process. Although this loss of methylation may be a result of changes in the chromatin structure, this is unlikely as the same chromatin changes occur in non-mammalian genomes in the absence of global demethylation (reviewed in Reik et al., 2001).

Several studies have attempted to identify the enzymes involved in demethylation in mammals. Two studies of oestrogen receptor-alpha responsive genes in cultured breast cancer cells noted cyclical patterns of demethylation and remethylation of CpG sites in target genes over a two hour period following treatment with oestrogen. Chromatin immunoprecipitation demonstrated coincident recruitment of DNMT3A and DNMT3B to the promoter regions of these genes, suggesting that these enzymes may be involved in the demethylation process. The authors suggest that the deaminase activity of these enzymes allows them to mediate oxidative deamination of methylcytosine in the absence of the

methyl donor AdoMet (S-adenosine L-methionine). This converts the methylcytosine to a thymine residue, leading to a guanine:thymine (G:T) mismatch, and subsequent removal of the thymine by thymine DNA glycosylase (TDG) and methyl-binding protein 4 (MBD4). Finally, this site is repaired by the base excision repair mechanisms, ultimately replacing the methylated cytosine residue with an unmethylated counterpart (Kangaspeska et al., 2008; Métivier et al., 2008; Ooi and Bestor, 2008). However, the short time frame of this reported demethylation would require DNMT3A and 3B to be extremely efficient deaminases. Furthermore, concentrations of AdoMet would have to be limiting *in vivo* in order for the conversion of methylcytosine to thymine to take place.

Demethylase activity has reportedly been observed in several studies of nuclear and whole cell extracts, including a protease-sensitive activity in murine erythroleukemia nuclear extracts and RNAse sensitive activity in whole cell rat myoblast extracts (Gjerset and Martin, 1982; Weiss et al., 1996). However, these reports have faced widespread criticism and subsequent characterisation of this demethylase activity has not been forthcoming. Thymine DNA glycosylase has also been identified as having methylcytosine glycosylase activity, however the efficiency of this reaction of too low for the enzyme to be a viable DNA demethylase (Cortázar et al., 2007; Jost, 1993).

It has also been proposed that methyl-binding protein 2 (MBD2) is able to catalyse direct removal of the methyl group from methylcytosine, however several other groups have subsequently been unable to repeat these experiments (Bhattacharya et al., 1999; Boeke et al., 2000; Ng et al., 1999; Wade et al., 1999). Additionally, *mbd2* knockout mice display similar levels and distribution of methylcytosine to wild type mice, alongside normal imprinting, suggesting

that this protein does not play an essential role in normal methylation patterning (Hendrich et al., 2001).

Finally, studies in both *Xenopus* and zebrafish have suggested that the nuclear protein Gadd45a may be involved in DNA demethylation via a base excision repair mechanism. *In vitro* experiments in HEK293T cells indicated that *GADD45A* overexpression lead to demethylation of a methylation-silenced reporter plasmid as well as global loss of DNA methylation. Additionally, Gadd45a is specifically recruited to the *Oct4* gene during active demethylation of this region in *Xenopus* oocytes (Barreto et al., 2007). Similar experiments in zebrafish revealed a demethylase activity on the introduction of *in vitro* methylated DNA into zebrafish embryos (Rai et al., 2008). This activity was shown to be dependent on the 5-methylcytosine deaminase AID, the G:T mismatch repair enzyme Mbd4 and Gadd45a. The authors suggest a demethylation mechanism similar to that described above, whereby methylcytosine is first deaminated to thymine by AID, followed by base excision repair to cytosine by Mbd4 and Gadd45a (Rai et al., 2008). However, another study of GADD45A activity using *in vitro* demethylation assays in human and mouse cell lines failed to reproduce these findings. Additionally, the authors note that Gadd45a is not expressed during early mouse development when active demethylation of the genome is observed, suggesting that this enzyme may not act as a DNA demethylase (Jin et al., 2008).

1.4.5.3 SAM availability and Hypomethylation

In addition to active demethylation or passive loss of methylation through decreased DNMT expression or activity, global hypomethylation may also result from changes in the availability of the substrates required by DNMTs. DNA methylation is achieved through an enzymatic reaction whereby a methyl group from the donor S-adenosylmethionine (SAM) is transferred to the C-5 in the pyrimidine ring of cytosine. In this process, SAM is converted to S-adenosylhomocysteine (SAH), which is in turn broken down to yield homocysteine and adenosine. Methionine is generated from homocysteine in a folate dependent reaction catalysed by MTR and MTRR and is finally converted to SAM to complete the cycle (reviewed in Laird and Jaenisch, 1996).

Disruption of SAM production, through altered metabolic activity or reduced substrate availability, could therefore lead to genome wide hypomethylation in cells. Since the conversion of homocysteine to SAM via methionine requires donation of a methyl group from substrates including dietary components such as folate, deficiency of these methyl donors could disrupt this process. A number of studies in rats support a causal relationship between methyl-deficient diets and decreased levels of SAM to global hypomethylation and cancer (reviewed in Szyf et al., 2004). Similarly, dietary methyl or folate deficiency together with high alcohol intake in humans have been suggested to increase risk of colorectal cancer, suggesting that altered SAM metabolism and subsequent disruption to DNA methylation may play an important role in the development of other cancers (Kim, 2005).

1.5 Polycomb Group Proteins and Target Gene Silencing

Polycomb group (PcG) proteins are epigenetic gene silencers first identified in *Drosophila* as developmental repressors of homeotic *Hox* genes (Ringrose and Paro, 2004). The vertebrate homologues of these genes play a similar role, acting to epigenetically silence gene expression through histone modification (Levine et al., 2002). Functionally, PcG proteins form conserved polycomb repressive complexes (PRCs), of which PRC2 plays a key role in normal and neoplastic development. PRC2 is composed of three core proteins – enhancer of Zeste homologue 2 *(EZH2),* suppressor of Zeste 12 *(SUZ12)* and enhancer of ectodermal development *(EED)* (Müller et al., 2002; Tie et al., 2001).

EZH2 catalyses the trimethylation of histone H3 lysine 27 (H3K27me3) and to a lesser extent trimethylation of histone H3 lysine 9 (H3K9me3) (Cao et al., 2002; Kuzmichev et al., 2002). This action of the PRC2 complex leads to gene repression by a number of downstream mechanisms. First, H3K27me3 is believed to provide a binding site for polycomb repressive complex 1 (PRC1), which in turn blocks the recruitment of transcriptional activation factors such as the ATP-dependent nucleosome remodelling complex *SWI/SNF* and prevents the initiation of transcription by RNA polymerase II (Dellino et al., 2004; Shao et al., 1999). Secondly, EZH2 is able to recruit DNA methyltransferases to certain target genes, resulting in gene silencing through promoter hypermethylation (Vire et al., 2006). Finally, there have been reported cases of PRC2 methylation of histone H1 lysine 26 (H1K26me), which may influence higher order chromatin structure by affecting the binding of heterochromatin binding protein 1 (HP1) (Daujat et al., 2005; Kuzmichev et al., 2004).

The closely related polycomb repressive complex 1 (PRC1) also plays a key role in the transcriptional silencing of polycomb target genes. The core components of PRC1 were originally identified in *Drosophila* as Polycomb (PC), Polyhomeotic (PH), Posterior Sex Combs (PSC) and Sex Combs Extra (SCE/RING), with the mammalian homologues of these proteins (CBX4, PH1, BMI1 and RING1B, respectively, in humans) also forming stable complexes (Shao et al., 1999; Simon and Kingston, 2009). PRC1 is able to bind the H3K27me3 chromatin modification through the chromodomain of PC, with localisation of the complex to target genes potentially mediated through this interaction (Min et al., 2003). Although the exact mechanisms of PRC1 dependent gene silencing have yet to be established, this action is believed to result from chromatin compaction and ubiquitylation of histone H2A by the complex (Wang et al., 2004).

1.5.1 Targeting of Polycomb Group Complexes

The mechanisms by which PRC1 and PRC2 are targeted to genes have been well studied in *Drosophila,* where PcG complexes are recruited to DNA sequences termed polycomb responsive elements (PREs) (Müller and Kassis, 2006). Although several PREs have been characterised through investigation of polycomb targets including the Hox genes and Engrailed, no consensus sequence between PREs has yet been identified. However, binding sites for the DNA binding protein Pleiohomeotic (PHO) and PHO-like are shared between PREs, with both proteins playing an essential role in PcG recruitment (Fritsch et al.,

1999; Simon and Kingston, 2009). Although mammalian PREs have yet to be defined, OCT4 and YY1 have been suggested to be involved in the recruitment of PcG complexes to target genes (Mihaly et al., 1998).

A number of investigations have also suggested a role for non-coding RNAs (ncRNAs) in PcG complex targeting. For example, the ncRNA *HOTAIR,* transcribed from the *HOXC* locus, has been shown to interact with PRC2 subunits. Furthermore, depletion of this RNA causes loss of H3K27me3 and desilencing of the *HOXD* locus in *trans* (Rinn et al., 2007). Similarly, the ncRNA *XIST* has been implicated in PRC2 targeting, with *in vitro* studies indicating that PRC2 interacts with a 28bp repeat element within the *XIST* transcript known as repA (Zhao et al., 2008). Further investigations will be required to establish the full role of ncRNAs and other mechanisms in PcG complex targeting in mammalian systems (Simon and Kingston, 2009).

1.5.2 The Role of PRC2

Gene repression by PRC2 is essential in normal development and cell fate decisions, specifically the maintenance of embryonic stem (ES) cell pluripotency. This role was first suggested by experiments which demonstrated that ES cell lines could not be established from *ezh2* deficient blastocysts (Erhardt et al., 2003). Further characterisation of PRC2 target genes by chromatin immunoprecipitation (ChIP) indicated that PcG proteins target a number of genes associated with differentiation and lineage specification (Boyer et al., 2006; Lee et al., 2006). Repression of these genes by PcG proteins is

therefore essential in maintaining an undifferentiated state in stem cells, as has been shown during neuronal differentiation (Bracken et al., 2006). Intriguingly, in ES cells many PcG target genes are co-occupied by OCT4, SOX2 and NANOG – transcription factors required for the maintenance of stem cell pluripotency (Sparmann and van Lohuizen, 2006). Although no direct interaction between these transcription factors and PcG proteins has been reported, it is possible that PcGs are recruited to these genes in order to maintain gene silencing and repress differentiation (Boyer et al., 2006; Lee et al., 2006). In support of this view, it has been observed that SUZ12 is displaced from certain target genes following OCT4 knockdown (Squazzo et al., 2006).

1.5.3 The Role of Bilvalent Chromatin in ES Cell Pluripotency

The repressive histone modification established by PRC2, H3K27me3, is one half of an unique chromatin state found at the promoter regions of key developmental genes along with methylated H3 lysine 4 (H3K4me) - a chromatin modification normally associated with active promoters. It is believed that this 'bivalent' chromatin domain maintains repression of developmental genes whilst keeping them poised for quick activation when growth signals are received (Bernstein et al., 2006). Genes associated with bivalent chromatin domains are transcribed at low levels in ES cells as a result of RNA polymerse binding, and become activated or repressed on the induction of differentiation through an increase in H3K4me or H3K27me3, respectively (Bernstein et al., 2006). The importance of these chromatin modifications in differentiation has been

confirmed by a genome wide study of chromatin modifications in mouse ES cells, neural protenitor cells and embryonic fibroblasts. This investigation confirmed that relative levels of H3K4 and H3K27 trimethylation effectively discriminates genes that are expressed, poised for expression, or stably repressed, therefore reflecting the current state and lineage potential of these cells (Mikkelsen et al., 2007). Furthermore, another study has shown that premature expression of genes associated with bivalent chromatin domains occurs in *Eed* deficient mouse ES cells (Azuara et al., 2006). Polycomb repressive complexes therefore play an important role in maintaining ES cell pluripotency through the creation and preservation of bivalent domains at key developmental genes.

1.5.4 PRC2 and Cancer

PRC2 also plays an important role in the development of cancer. *EZH2* up-regulation was initially identified in prostate cancer and has subsequently been shown in a variety of neoplasms including several types of lymphoma, breast cancer and melanoma (Varambally et al., 2002). Additionally, *SUZ12* shows similar overexpression in colon and breast cancer (Sparmann and van Lohuizen, 2006). Further characterisation has demonstrated that *EZH2* has many properties consistent with an oncogene. In particular, *EZH2* overexpression increases cell proliferation, invasion of benign cells, colony formation and *in vivo* tumour growth in nude mice (Bracken et al., 2003; Kleer et al., 2003; Varambally et al., 2002). Conversely, *EZH2* knockdown results in growth arrest of cancer cells, as well as decreasing tumour growth and metastasis *in vivo* (Croonquist and Van,

2005; Varambally et al., 2002). Finally, high levels of *EZH2* expression are associated with metastatic states in both prostate and breast cancer (Croonquist and Van, 2005; Kleer et al., 2003; Varambally et al., 2002).

The repressive action of PRC2 suggests a role for this complex in gene silencing during tumourigenesis (Schlesinger et al., 2007). In support of this view, a study of DNA methylation in colorectal cancers and matched normal mucosa indicated that PcG targets are up to 12 fold more likely to be hypermethylated than non-targets in the cancer (Widschwendter et al., 2007). Additionally, there is a strong inverse correlation between PRC2 occupancy and RNA polymerase II occupancy at target gene promoters, suggesting that these genes are not transcribed (Lee et al., 2006; Squazzo et al., 2006). However, other studies have also suggested a role for PRC2 in gene activation. In particular, PRC2 knockdown in colon cancer cells resulted in down-regulation of a small subset of genes and *EZH2* has also been linked to activation of cell cycle control genes in breast cancer cells (Simon and Lange, 2008).

There are two possible ways in which this gene repression could contribute towards transformation. Firstly, PcG complexes may silence tumour suppressor genes, which normally act to limit uncontrolled cellular proliferation. Secondly, the involvement of PcG proteins in maintaining pluripotency may contribute to the multipotency of cancer cells (Sparmann and van Lohuizen, 2006). Shared properties between cancer cells and stem cells, namely the ability to divide indefinitely and generate a diverse set of progeny, suggest that tumourigenesis may take place in, or involve a transition towards undifferentiated or de-differentiated cells. The role of *EZH2* in maintaining pluripotency by repressing differentiation suggests that the PRC2 complex may

play a similar role during transformation (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). Furthermore, similarities in gene expression signatures between poorly differentiated tumours and ES cells comprise a number of PRC2 target genes alongside OCT4, SOX2, NANOG and c-MYC target genes, supporting the view that aberrant PRC2 activity plays an important role in maintaining transformed cells in a more stem cell-like state (Ben-Porath et al., 2008).

As well as increased expression of PcG proteins in cancer, there are indications that PRC2 target gene preference may also be altered in cancer. Although PcG proteins normally target developmental transcription factors in embryonic cell lines, SUZ12 has been shown to also target glycoprotein genes in adult tumour cells (Squazzo et al., 2006). Alterations to the composition and substrate specificity of PcG complexes may account for these changes in target preference. EZH2 is able to associate with different EED isoforms, of which there are four. Classical PRC2 contains the longest form of EED, EED1, whereas shorter isoforms associate with EZH2 to form alternative 'PRC2-like' complexes PRC3 and PRC4. PRC4 is selectively enriched in cancer cells and formation of this complex is favoured when EZH2 is overexpressed (Kuzmichev et al., 2005). PRC4 displays a different substrate affinity to PRC2 – preferring to methylate histone H1K26 rather than H3K27. Furthermore, the histone deacetylase SIRT1 also forms part of the complex and targets H1K26, potentially allowing the complex to deacetylate this histone before a methylation event (Kuzmichev et al., 2004). However, a study of H1 methylation by recombinant EZH2 complexes in the presence of different EED isoforms failed to detect any significant variation in substrate preference (Martin et al., 2006).

1.5.5 Mechanisms of PRC2 Up-regulation

Several studies have investigated the mechanisms by which expression of PRC2 components are regulated (Figure 1.1). A study of *EZH2* regulation identified microRNA-101 as a negative regulator of its expression and function. Furthermore, the genomic region containing this micro-RNA was deleted in some prostate tumours, leading to *EZH2* overexpression (Varambally et al., 2008).

A link between the control of EZH2 expression and the oncogene *MYC* has also been identified. Utilizing a murine model of lymphoma and human Burkitt's lymphoma samples, MYC was shown to be a negative regulator of miR-26a, which is in turn an inhibitor of *EZH2* expression. Increased *MYC* expression or activity in cancer can therefore lead to concurrent increase in EZH2 expression via this microRNA mediated mechanism (Sander et al., 2008).

EZH2, *SUZ12* and *EED* have been shown to be downstream of the pRb-E2F pathway, with the E2F transcription factor binding to and activating all three promoters in normal and transformed cells (Bracken et al., 2003; Müller et al., 2001; Weinmann et al., 2001).

Finally, viral oncogenes have also been shown to regulate expression of PRC2 components. For example, human papillomavirus 16 (HPV16) E7 has been shown to increase EZH2 expression by releasing E2F. Additionally, a separate study has demonstrated that *EZH2* expression is required for proliferation of HPV positive tumour cells and contributes to the apoptotic resistance of cervical cancer cells (Holland et al., 2008).

Figure 1.1

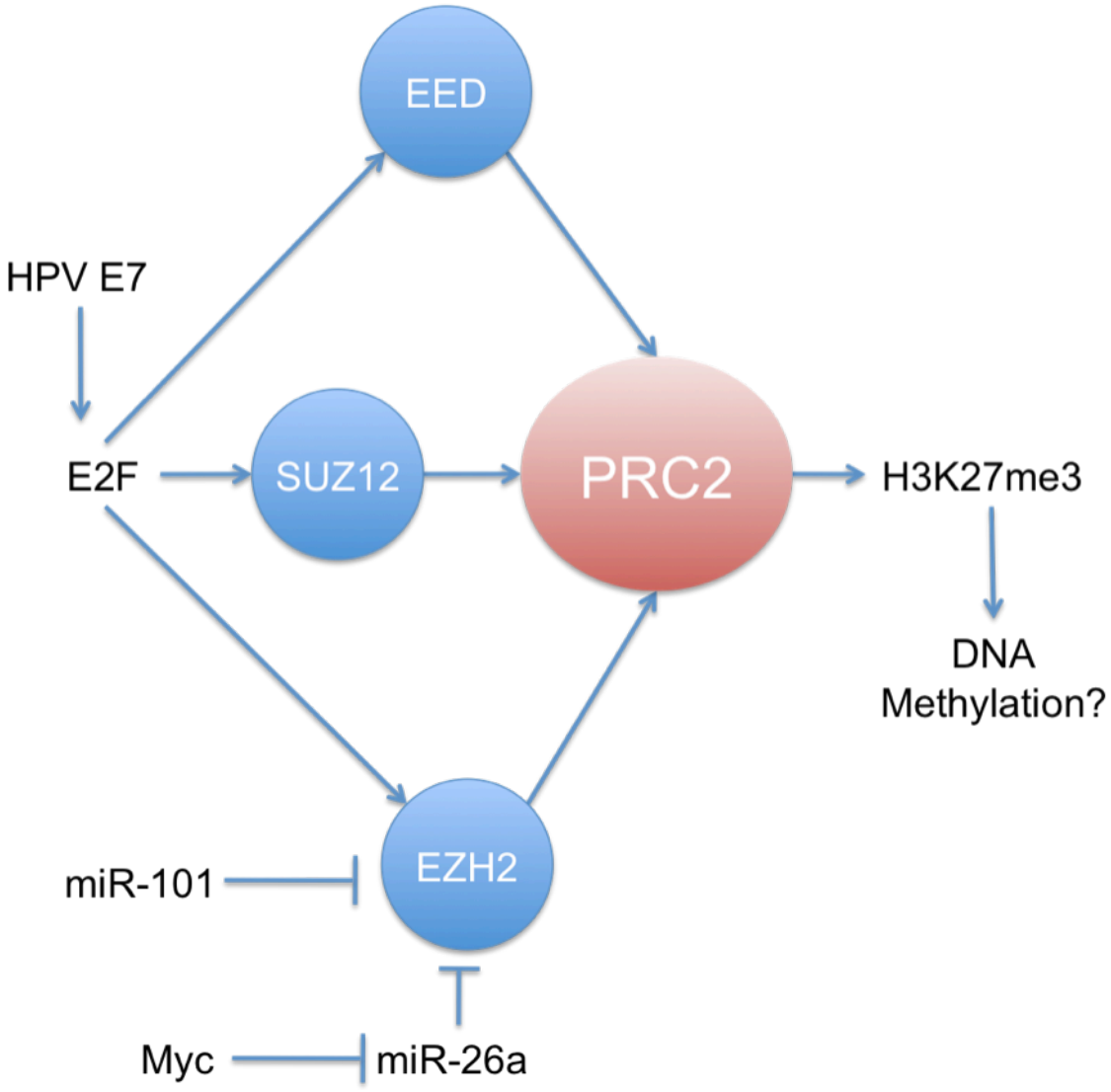


Figure 1.1: Summary of PRC2 regulation. Expression of the PRC2 components EZH2, EED and SUZ12 has been shown to be governed by a number of mechanisms, including direct transcriptional regulation of all components by the transcription factor E2F and negative regulation of EZH2 expression and function by microRNAs.

1.6 The Cancer Stem Cell Theory

Identifying the 'normal' tumour originating cells which are subject to oncogenic changes has been a subject of much debate in recent years. A number of observations have questioned the traditional view of terminally differentiated cells acquiring oncogenic hits. In particular, the heterogeneous nature of tumours, comprising a wide variety of cells in different states of maturity and transformation, has been difficult to account for. Furthermore, only a small proportion of cancer cell lines are able to form tumours in experimental models, often only with the introduction of a large quantity of cells. Finally, cancer cell lines and tumour xenografts can only be obtained from patient tumours with a very low efficiency (Bjerkvig et al., 2005; Polyak and Hahn, 2006; Reya et al., 2001).

Evidence for the existence of cancer stem cells first arose during an *in vivo* study of acute myeloid leukaemia (AML) whereby it was demonstrated that only a small subset (1 in 250,000) of cancer cells were able to recapitulate a full cancer when transplanted into immunodeficient mice, with the $CD34^{+}/CD38^{-}$ fraction of cells enriched for tumour forming ability (Lapidot et al., 1994). Cancer stem cells have subsequently been identified in a variety of different cancers (Visvader and Lindeman, 2008). The first report of cancer stem cells in a solid tumour came from a study of breast cancer in which cells from 9 patients were isolated and sorted on the basis of surface marker expression. As few as 200 cells from the fraction of cells expressing CD44 and low levels of CD24 ($CD44^{+}/CD24^{-/low}$) were required to form tumours when injected into mice,

whereas injection of 100-fold more cells from the $CD44^+/CD24^+$ fraction failed to form tumours (Al-Hajj et al., 2003). Cancer stem cells have also been isolated using CD133 in a variety of cancers including several types of brain tumour, as well as colorectal and pancreatic carcinoma (Beier et al., 2007; Hermann et al., 2007; O'Brien et al., 2007; Ricci-Vitiani et al., 2007; Singh et al., 2004; Taylor et al., 2005).

1.6.1 Similarities Between Normal and Cancer Stem Cells

1.6.1.1 Signaling Pathways

Many of the signalling pathways involved in the self-renewal capacity of normal stem cells are also implicated in development of cancer. In particular, the Wnt, Notch, Hedgehog, PTEN and BMI1 pathways all play an important role in maintaining self-renewal in normal stem cells and have been shown to promote neoplastic proliferation when deregulated (Pardal et al., 2003). For example, Wnt signalling plays in important role in regulating intestinal stem cell self renewal, in part by controlling cell migration and proliferation through downstream ephrin family proteins and c-MYC respectively (Dodge et al., 2005; van de Wetering et al., 2002). Mutations that activate Wnt signalling in intestinal progenitors lead to hyperproliferation of these cells and subsequent development of benign polyps in which multi-lineage differentiation can be observed (Moser et al., 1992; Powell et al., 1992; van de Wetering et al., 2002). Wnt signalling has also been implicated in the maintenance of haematopoietic stem cell self renewal as well

as the renewal of CML stem cells *in vivo* (Austin et al., 1997; Reya and Clevers, 2005; Van Den Berg et al., 1998; Zhao et al., 2007).

Notch signalling has also been shown to play an important role in haematopoietic stem cell maintenance, with stimulation by the Jagged-1 ligand increasing progenitor activity of these cell *in vitro* and *in vivo* (Karanu et al., 2000; Varnum-Finney et al., 2000). Increased self-renewal of haematopoietic stem cells has also been observed *in vitro* in response to sonic hedgehog signalling in combination with other growth factors (Bhardwaj et al., 2001).

1.6.1.2 Polycomb Group Activity

Polycomb group proteins also play an important role in the maintenance of stem cells through their repressive action. A particularly important protein in this group is BMI1, which promotes cell proliferation through inhibition of CDKN2A, encoding for the cyclin dependent kinase inhibitors INK4A (*p16*) and ARF (*p14*) (Jacobs et al., 1999a; Jacobs et al., 1999b). Haematopoietic stem cells lacking BMI1 show limited self renewal capacity (Park et al., 2003). Additionally, proliferation of leukaemic stem cell has been shown to be dependent on BMI1, with deficient stem cells lacking the ability to recapitulate the cancer when transplanted into immunodeficient mice (Lessard and Sauvageau, 2003). The dependence of both normal and cancer stem cells on similar signalling pathways is a further indication that cancer stem cells may arise as a result of dysregulation of self-renewal pathways in normal stem cells.

1.6.2 Normal Cells Subject To Transformation

The identity of the cells from which cancer stem cells originate has been a subject of widespread debate. The self-renewal properties of cancer stem cells suggest that they develop either from normal adult stem cells, through the acquisition of oncogenic mutations, or from differentiated cells that 'de-differentiate' into a more stem-cell like state (Zilberman, 2007). The fact that several mutations or epigenetic changes are required for transformation to take place suggests that cells with a longer lifespan, such as stem cells, are more likely to accumulate such oncogenic changes (Knudson et al., 1973). Since the turnover of differentiated cells in many organs is high, it has been suggested that the adult stem cells in these organs are subject to transformation as these are the only cells with a long enough life span. Studies of leukaemic stem cells have demonstrated that they display similar surface markers to haematopoietic stem cells, suggesting that they may arise from transformation of these progenitors (Bonnet and Dick, 1997). In support of this hypothesis, leukaemogenic mutations increase the proliferation rate and block differentiation of haematopoietic stem cells (Kelly and Gilliland, 2002; Pereira et al., 1998). In addition, there are numerous phenotypic similarities between soft tissue sarcoma cells and mesenchymal progenitor cells, and these progenitors can be converted into Ewing's sarcoma through expression of the fusion protein ETS-FL1 (Riggi et al., 2005). Additionally, a mouse model in which the SYT-SSX fusion protein is conditionally expressed has been used to show that synovial sarcomas derive from myogenic precursor cells (Haldar et al., 2007). Although in both these cases the progenitor cells are derived from the same tissue of origin as the cancer,

gastric cancers can originate from transformation of bone marrow derived progenitor cells (Houghton et al., 2004). These studies suggest that cancers arise following oncogenic changes in adult stem cells or progenitor cells.

The alternate hypothesis is that cancer stem cells could originate from committed progenitors or fully differentiated cells that acquire self-renewal capacities and de-differentiate towards a more stem cell like state. In line with this, a number of fusion proteins including MOZ-TIF2 and MLL-AF9 have been shown to activate self-renewal systems and confer stem cell like properties on committed haematopoietic progenitor cells, leading to the generation of cancer stem cells (Huntly et al., 2004; Krivtsov et al., 2006). Similarly, a β-catenin mutation promotes self-renewal in granulocyte-macrophage progenitors in patients with chronic myeloid leukaemia (Jamieson et al., 2004). A number studies have demonstrated that pluripotent cells can be created from fully differentiated cells through the expression of a small number of transcription factors – namely Oct4, Sox2, Klf4 and c-Myc, although some of these factors have been shown to be dispensable or can be replaced with others (Feng et al., 2009; Nakagawa et al., 2007; Takahashi et al., 2007). These induced pluripotent cells (iPS) are virtually indistinguishable from normal embryonic stem cells and are able to differentiate into cell types from the three germ layers (Takahashi et al., 2007). It remains a strong possibility, therefore, that similar 'de-differentiation' events occur in differentiated cells during the development of cancer, allowing the cells to acquire more stem-cell like characteristics.

1.6.3 Limitations Of The Cancer Stem Cell Theory

1.6.3.1 Differences In Surface Markers

A number of cell surface markers and dye efflux assays have been described for the selective isolation of cancer stem cells from a variety of cancers. These include CD133, CD24, CD44, EpCAM, THY1 as well as exclusion of $Hoechst_{33342}$ dye from a side population (Visvader and Lindeman, 2008). In many cases, cancer stem cells share many of the surface markers of normal stem cells from the same organ. Although markers such as CD133 and CD44 are present on cancer stem cells from a variety of origins, a single definitive marker for cancer stem cells has not yet been identified. In fact, there is a great degree of heterogeneity in the cancer stem cell markers reported by different groups for the same tumour type. For example, in breast cancer, high aldehyde dehydrogenase 1 (ALDH1) activity has been reported in the tumourigenic fraction of cells. However, very little overlap was observed between these cells and the $CD44^+/CD24^-$ phenotype previously described for breast cancer stem cells (Al-Hajj et al., 2003; Ginestier et al., 2007). There also appears to be a degree of heterogeneity within individual cancer stem cell populations. A study of glioblastoma cancer stem cells indicated that both $CD133^+$ and $CD133^-$ cells could form tumours when injected into nude mice (Beier et al., 2007). Another investigation of cell lines derived from *brca1* deficient mouse mammary tumours showed that either $CD44^+/CD42^-$ or $CD133^+$ cells could form tumours when injected into immunodeficient mice in small numbers, with no overlap between the two populations (Wright et al., 2008).

As well as differences between the reported markers for cancer stem cells, there is also a great degree of variation in reports of the proportion of tumour cells with tumour initiating potential. The proportion of $ABCB5^+$ cells reported in melanomas varies from 1.6-20% of total tumour cells and the proportion of $CD133^+$ cells in colorectal carcinomas ranges from 1.8-24.5% (O'Brien et al., 2007; Schatton et al., 2008). These findings indicate that although considerable progress has been made in the identification and characterization of cancer stem cells, there is significant variation in the phenotype and percentage of putative cancer stem cells in different tumours. Further studies are required to determine the source and significance of this variation.

1.6.3.2 Differences In Transplantation Efficiency

Although serially injecting cells into immunodeficient mice is currently the gold standard test for cancer stem cells, this method has a number of drawbacks. First, differences between both species and tissue specific microenvironments may introduce a selection bias for cells with enhanced immune evasion and propagation potential rather than genuine cancer stem cells (Kelly et al., 2007). A number of modifications to the standard protocols have been trialled in order to minimise such differences and recapitulate a more permissive environment for the transplanted cells. For example, human breast fibroblasts can be injected into a mouse mammary pad in order to create a more 'humanized' environment for xenotransplantation. Although transplanted normal breast stem cells are able to grow at these sites, repopulation frequency remains low (Kuperwasser et al.,

2004). Alternative strategies include mixing the stem cells with other cancer associated cells such as fibroblasts or mesenchymal stem cells may help to recreate a more appropriate niche (Karnoub et al., 2007).

As well as the immediate cellular environment, other systemic factors may also play a role in the efficiency of tumour formation in mice. For example, oestrogen levels were shown to be important in the growth of oestrogen-receptor negative breast cancers in mice, leading to recruitment of stromal cells important for the development of these tumours (Gupta et al., 2007).

The efficiency of tumour formation in mice can also be affected by residual immune function in the host. Although the most commonly used non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice are severely immunocompromised, these mice do retain a small degree of natural killer cell activity. This could potentially result in an underestimation of the efficiency at which transplanted cells are able to recapitulate tumours. In fact, a study utilizing the more severely immunocompromised NOD/SCID interleukin-2 receptor gamma chain null ($il2rg^{-/-}$) mice demonstrated that the proportion of detectable tumourigenic melanoma cells can be drastically increased (Quintana et al., 2008). Injection of unselected melanoma cells from 12 different patients into these mice showed that approximately 25% of these cells were able to form tumours. Furthermore, in single cell transplants, 27% of unselected melanoma cells were able to form tumours. Finally, a screen of surface marker expression in these tumour forming cells failed to detect any suitable markers to distinguish these cells from the population unable to form tumours. These results show that the proportion of cancer initiating cells in melanoma is much higher than was previously reported and that the tumour forming cells are apparently

phenotypically heterogeneous and lack the classic cancer stem cell markers previously reported (Quintana et al., 2008).

Finally, the site of transplantation may also have an effect on the frequency of tumour initiation. A number of studies have reported significant differences in tumour initiating potential when cells are injected into different sites in mice. A study of glioblastoma derived cell transplantation revealed an increase in tumour initiation frequency from 50% when injected subcutaneously to 100% when cells were injected into the cranium (Galli et al., 2004). However, strict orthotopic transplantation of the cells does not always seem necessary – injection of colorectal cancer stem cells under the kidney capsule can also increase tumour formation to 100% compared to 30% subcutaneously (Visvader and Lindeman, 2008). This suggests that a number of factors – such as differences in vascularisation, growth factors and extracellular matrix at the site of injection - play a role in generating the observed differences in tumour formation efficiency.

1.7 The Epigenetic Progenitor Origin of Cancer

A landmark paper by Feinberg *et al.* has attempted to unify the notions of genetic and epigenetic changes alongside the cancer stem cell theory by exploring the role of progenitor cells in cancer formation and proposing a step-wise mechanistic basis for epigenetic disruption and transformation (Feinberg et al., 2006).

In this report, the authors note that many lines of evidence indicate that stem cells or progenitor cells may be the targets of transformation. Furthermore, epigenetic mechanisms are key in maintaining this stem cell identity, and are also commonly dysregulated during the development of cancer. Taken together, they postulate that early epigenetic disruption of stem cells is a key determinant of cancer risk, tumour progression and heterogeneity. Five key observations are provided to support the notion that cancers arise from epigenetically disrupted progenitor cells: firstly, *in vitro* studies have demonstrated that tumour related growth properties of both leukaemia and solid tumours are reversible, indicating an epigenetic basis to this phenotype (Lotem and Sachs, 2002; Sachs, 1986; Yuspa, 1983). Secondly, the ubiquity of global epigenetic changes such as genome wide hypomethylation and promoter hypermethylation in cancers, including benign neoplasms, indicates that these changes are likely to occur before any genetic changes that lead to transformation (Feinberg and Vogelstein, 1983; Goelz et al., 1985). Thirdly, a study has demonstrated that mouse melanoma nuclei can be used to clone a normal mouse, indicating that many of the tumourigenic properties of these cells are epigenetically controlled and can be reversed (Hochedlinger et al., 2004). Fourthly, serial transplantation of

selected cell populations derived from tumours has demonstrated that a small subset of tumour cells retain self-renewal capabilities and are able to recapitulate a full tumour with a range of cell types (Kim et al., 2005; Pardal et al., 2003; Singh et al., 2004). Finally, a study of colorectal cancer cases associated with loss of imprinting (LOI) of the IGF2 gene demonstrated that this LOI is observed in the apparently normal colonic epithelium of these patients, and is therefore likely to have occurred before further tumourigenic hits. Additionally, LOI of IGF2 in both humans and mice is associated with a shift in the proportion of progenitor and differentiated cells in the colonic epithelium (Cui et al., 2003; Sakatani et al., 2005; Woodson et al., 2004).

In this model, tumourigenesis occurs in three distinct stages. The first of these involves epigenetic disruption of tumour progenitor genes – genes that can increase cancer likelihood by increasing cell 'stemness', or properties of self-renewal and pluripotency over a tendency towards limited replicative potential and differentiation. Candidate tumour progenitor genes include *OCT4*, *IGF2,* E-cadherin, *FOXD3* and *EZH2*, all genes involved in the maintenance of stem cell pluripotency. Following this epigenetic disruption in a progenitor cell population, the second step towards transformation can take place – an initiating mutation that is specific to the cancer type. For example, specific mutations of *APC* occur in colorectal cancer, and specific recombination events lead to the formation of fusion proteins such as BCR-ABL in CML. Finally, genetic and epigenetic plasticity is required for further tumourigenic changes to take place. This can take the form of genomic instability or changes to DNA methylation and histone modifications, or to the mechanisms governing these epigenetic marks.

1.8 The Stem Cell Model

In order to investigate the potential link between adult stem cells and cancer, Funes *et al.* have established a line of primary mesenchymal stem cells - thought to be the origin of various types of sarcoma - in which five oncogenic steps towards a fully transformed state are sequentially introduced (Figure 1.2a; Funes et al., 2007).

Primary human MSC were isolated from the bone marrow of a healthy 34 year old male donor by density gradient centrifugation, with MSC colonies formed after plating designated passage 0 and subsequently passaged approximately 7 days after isolation. Isolated MSC were shown by FACS to express the MSC markers including SH2, SH4 and Stro-1 and did not express negative markers of MSC including CD34 (haematopoietic stem cell maker) and CD45 (pan-leukocycte marker). Isolated MSC were also shown to be able to differentiate into into adipocytes, chrondrocytes and osteocytes, confirming the multipotency of these cells (Funes et al., 2007).

The primary MSC were transformed in a step-wise manner through the sequential expression of the following genes:

- hTERT (the catalytic subunit of human telomerase) - necessary to extend the life span of MSC in culture
- HPV-16 E6 and E7 to abrogate the functions of p53 and pRb tumour suppressor family members respectively
- SV40 small T antigen (ST) to inactivate protein phosphatase 2A (PP2A), resulting in stabilization of the c-Myc oncogene

- Constitutively active Ras (H-Ras^{V12})

These genetic elements were sequentially introduced into the cells by retroviral infection, with successful mRNA and protein expression confirmed by qRT-PCR and western blot respectively (Figure 1.2b). Previous models of step-wise transformation have suggested that although the total number of oncogenic changes required for transformation of a given cell type remains fixed, the order in which these events take place can be flexible (Fearon and Vogelstein, 1990; Hahn et al., 1999; Rangarajan et al., 2004). The particular sequence utilized in the MSC model was chosen for a number of reasons. Firstly, initial expression of hTERT was necessary to extend the lifespan of primary MSC in culture by overcoming replicative senescence (Counter et al., 1992; Hayflick, 1965; Hayflick and Moorhead, 1961). Secondly, inactivation of tumour suppressor genes was necessary before the expression of oncogenes in order to avoid oncogene induced senescence. Although the relative order of c-Myc and H-Ras^{V12} over-expression (at MSC 4 and MSC 5 in the model) were shown to be interchangeable, it remains possible that variability also exists in the order in which the earlier 'hits' are introduced - for example, SV40 small-T antigen could have been a suitable first 'hit' in place of hTERT (Funes et al., 2007). However, the order of genetic aberrations utilized in this model allows specific changes linked to immortalization (through hTERT expression) to be distinguished from the other transforming effects of viral oncoproteins such as SV40 small-T antigen.

The validity of this model was demonstrated in a number of ways. In a soft agarose test of anchorage independent growth, cells transduced with hTERT,

E6, E7 and ST formed small colonies in soft agarose after 12 days, with the final addition of H-Ras resulting in rapidly growing large colonies (Figure 1.2c). Furthermore, cells with all 5 transforming steps were able to induce tumours when injected into immunodeficient mice (Figure 1.2d). Gene expression microarray analysis of these cells also showed typical hallmark changes to transformed cells: up-regulation of anti-apoptotic pathways, up-regulation of angiogenic genes and down-regulation of MHC class I and II molecules. Finally, gene expression microarrays also show that the expression of genes downstream of pRB, p53, PP2A and H-Ras are affected with each transformation step.

This system provides a well defined model of transformation in which known genetic hits can be linked to downstream genetic or epigenetic changes. Furthermore, this model provides a unique insight into processes occurring during transformation in pre-malignant cells.

This stepwise model of transformation has been used to investigate cancer associated changes in bioenergetics and to identify novel putative anti-cancer drug targets (Flanagan et al., 2009; Funes et al., 2007). I hypothesise that epigenetic changes traditionally associated with cancer cells will also occur in this stem cell model. By providing a well defined background in which known genetic hits may be linked to observed epigenetic changes, the model could be useful to investigate the epigenetic hallmarks of cancer.

Figure 1.2

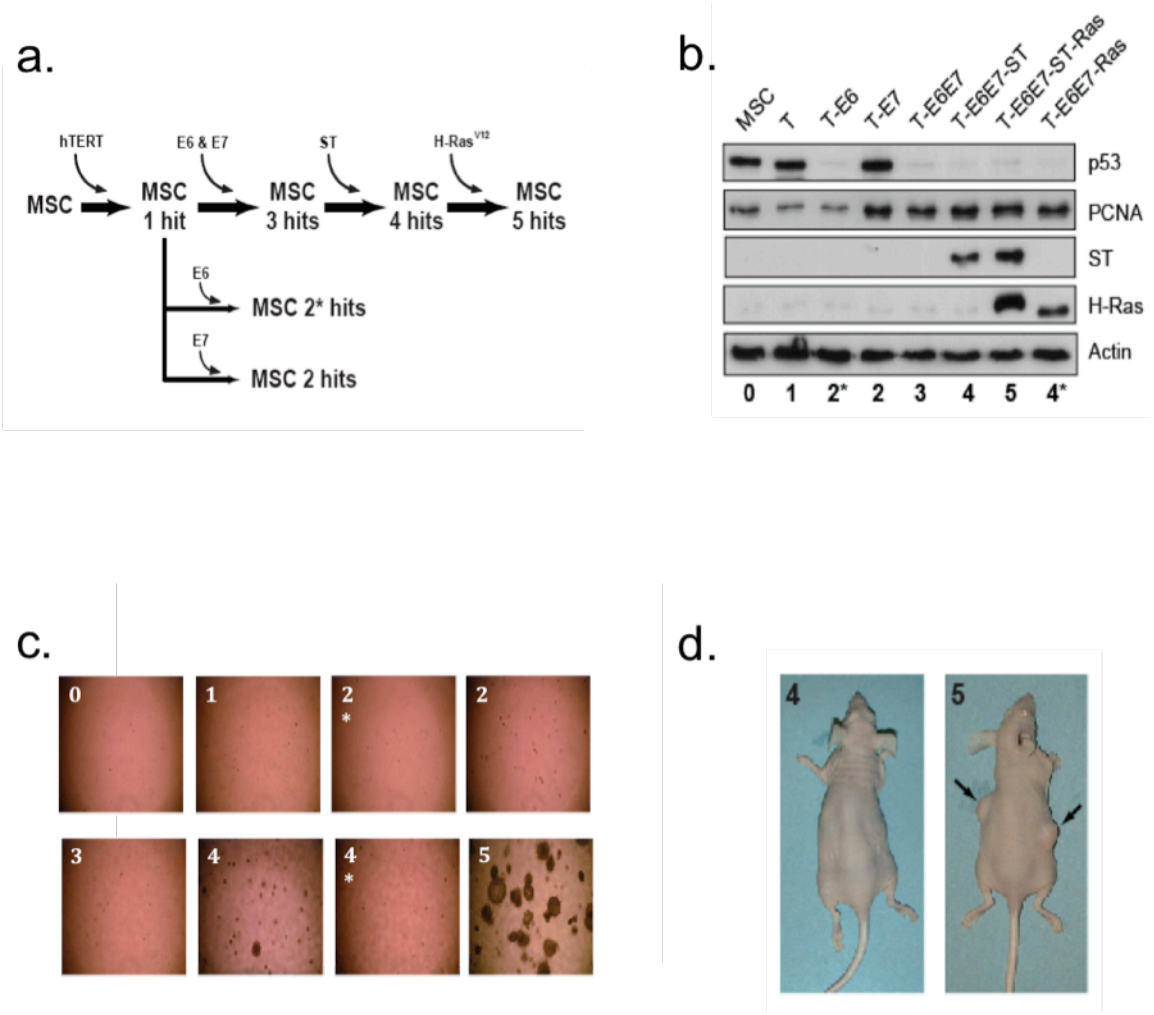


Figure 1.2: Step-wise transformation of human MSC. (a) Schematic diagram of stepwise MSC transformation. **(b)** Cell lines were probed by antibodies against the expressed oncoproteins or their downstream cellular targets. Proliferating cell nuclear antigen (PCNA) reflects cell entry in S phase after pRB inactivation by E7. **(c)** Anchorage independent growth of MSC infected with empty retroviruses or combinations of different oncogenes. **(d)** Tumours formed in mice inoculated with 5 hit MSC (arrows) but not in those inoculated with 4 hit MSC (left). Adapted from Funes *et al.*, 2007.

1.9 Aims of this thesis

The aim of this work is to investigate epigenetic changes that occur during stepwise transformation in this MSC model. This research centres on two major themes – investigation of global hypomethylation and of histone modifications mediated by the PRC2 complex. Firstly, I aim to determine whether global hypomethylation occurs in this stepwise model of transformation, and establish at which point this takes place. Furthermore, if genome-wide hypomethylation is observed, I will explore the causes of this loss of methylation. Secondly, an investigation of the role of PRC2 in this step-wise transformation will be conducted. In particular, I aim to establish whether the three core components of PRC2 (EZH2, EED and SUZ12) are up-regulated during stepwise transformation and whether PRC2 target genes are down-regulated through the repressive action of this complex. If this is the case, I aim to explore whether PRC2 activity is required for transformation to take place in this model.

Chapter 2: Materials and Methods

2.1 Cell Culture Techniques

2.1.1 Cell Lines and Maintenance

Mesenchymal stem cells (MSC) and human fibroblast cells (HF) were transformed in a step-wise manner through the sequential addition of the following genetic elements (cell line abbreviations are denoted in bold) (Funes et al., 2007):

1. hTERT (the catalytic subunit of human telomerase) - necessary to extend the life span of MSC and HF in culture **(MSC 0 / HF 0)**
2. a. HPV-16 E6 to abrogate the function of p53 **(MSC 2E6)**
   b. HPV-16 E7 to inhibit the function of pRb **(MSC 2E7)**
3. A combination of HPV-16 E6 and E7 to inhibit both p53 and pRb tumour suppressor pathways **(MSC 3 / HF 3)**
4. SV40 small T antigen (ST) to inactivate protein phosphatase 2a (PP2a), resulting in stabilization of the c-Myc oncogene **(MSC 4 / HF 4)**
5. Constitutively active H-Ras^{V12} **(MSC 5 / HF 5)**

Mesenchymal stem cells (MSC) were cultured in Mesencult medium supplemented with 10% human serum (StemCell Technologies, Vancouver, Canada) and 1 ng/ml basic fibroblast growth factor (R&D Systems, Minneapolis, MN).

Human fibroblasts (HF) and Human Embryonic Kidney 293T cells were maintained in Dulbecco's Modified Eagles medium (DMEM) (Gibco, Invitrogen, Carlsbad, USA) supplemented with 10% heat inactivated foetal calf serum (FCS; Gibco), 100 units/ml penicillin G, and 100 μg/ml streptomycin (Gibco).

All cells were grown at 37°C, 5% CO_2 and 100% humidity. Cells were kept at subconfluent levels and were typically passaged when 70-80% confluent. The cells were washed with 4ml Dulbecco's phosphate buffered saline (DPBS) (Sigma Aldrich, St. Louis, US) and incubated with 1ml 0.05% trypsin EDTA (Gibco) for 3-5 minutes at 37°C. An equal volume of full media was added to quench the reaction and the cells were centrifuged at 1200rpm for 5 minutes. The cell pellet was resuspended in an appropriate volume of fresh media and seeded at the following dilutions (Table 2.1).

Table 2.1: Typical seeding dilutions for cell culture

Cell Line	Dilution
MSC 0 / HF 0	1:3
MSC 1 / HF 1	1:3
MSC 2E6	1:5
MSC 2E7	1:7
MSC 3 / HF 3	1:10
MSC 4 / HF 4	1:12
MSC 5 / HF 5	1:12
293T	1:6

2.1.2 Cryogenic Storage and Recovery of Cells

Cells were cryogenically stored by resuspending the live cell pellet in foetal bovine serum (FBS) supplemented with 10% dimethyl sulphoxide (DMSO) to a concentration of 10^5 cells per millilitre. The cell suspension was aliquotted into polypropylene cryotubes (Corning, New York, US) and placed at -80°C overnight in a cryogenic storage vessel containing 250ml isopropanol to slow the freezing process. Once frozen, tubes were stored in liquid nitrogen until used.

Frozen cells were thawed at 37°C before the cell suspension was added to 5ml full media. Cells were centrifuged at 1200rpm for 5 minutes and the cell pellet was resuspended in full media and seeded into a 10cm dish.

2.1.3 Cell Harvesting for DNA and RNA Extraction

Cells were washed with 4ml Dulbecco's phosphate buffered saline (DPBS) (Sigma) and incubated with 1ml trypsin (Gibco) for 3-5 minutes at 37°C. Once detached, the trypsin was inactivated by the addition of 4ml fresh media, and the cells were pelleted in a 15ml Falcon tube by centrifugation at 1200rpm for 5 minutes. The cells were washed with 1ml DPBS and pelleted in a 1.5ml Eppendorf tube by centrifugation at 2000rpm for 5 minutes at 4°C. Cell pellets were stored at -80°C before use.

2.1.4 Clonal Cell Population Isolation

Clone isolation was performed by seeding cells at a density of 100 cells per 10cm dish and incubating at 37°C until colonies were visible. Individual clones were removed by placing a cloning cylinder (Sigma Aldrich, St. Louis, US) around the cells and adding 100µl trypsin 0.05% EDTA. The trypsin was neutralized by addition of an equal volume of full media, and the cells were transferred to a 96 well plate with the media replaced once the cells had attached. Cells were grown as normal and expanded until the required number were available.

2.1.5 5-Azacytidine treatment

Cell culture media was supplemented with 5-aza-2'-deoxycytidine (Sigma) to a final concentration of 5µM and added to cells 24 hours after seeding. Fresh drug-containing media was added to the cells every 24 hours, with drug treatment continuing for a total of four days, after which time cells were harvested for DNA or RNA extraction.

2.1.6 *Cell Viability and Growth Assay*

Cell viability and growth were assessed using the CellTiter 96 AQueous One Solution Cell Proliferation Assay (Promega, Madison, US). This assay utilizes the compound 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt (MTS) that is reduced to a coloured

formazan product in a reaction dependent on NADPH or NADH, levels of which are in turn directly proportional to the number of metabolically active cells present in a culture.

Cells were seeded at a density of 1000 cells per well containing 100µl media in a 96 well plate. Viable cell numbers were determined at the initial time point and every 24 hours by adding 20µl of the detection reagent to each well and incubating at 37°C for 2 hours. Absorbance at 490nm was recorded using a Fluoroskan spectrophotometer (Thermo Fisher, Waltham, USA). This measurement is directly proportional to the number of live cells in each well. Cell lines were assayed in triplicate for each time point, with blank readings obtained from wells containing media plus reagent only.

2.1.7 Soft Agarose Transformation Assay

Transformation was assessed by anchorage independent growth of cells in soft agarose. 6 well plates were prepared with a 2ml layer of 0.6% low melting point agarose (Sigma) in DMEM (Gibco). 10^4 cells were suspended in 0.35% low melting point agarose in DMEM and seeded in triplicate on the 0.6% agarose layer. Cells were grown at 37°C for 12 days, with colony number and size quantified visually at 40x magnification on the Axiovert 100 microscope (Carl Zeiss, Oberkochen, Germany) and with the SynGene G:Box and GeneTools software (Syngene, Frederick, USA).

2.2 Transfection and Virus Production

2.2.1 Cell Transfection

Cells were transfected using the Fugene transfection reagent (Roche, Basel, Switzerland) according to the manufacturer's instructions. To transfect one well of a 6 well plate, 1µg of DNA and 3µl of Fugene (3x the amount of DNA) were added to two separate 1.5ml Eppendorf tubes containing 50µl of serum free Optimem (Gibco) and incubated at room temperature (22°C) for 5 minutes. The contents of the two tubes were combined to a total volume of 100µl and incubated at room temperature (22°C) for a further 15 minutes. During this time, the target cells were washed with PBS and media replaced with serum free Optimem. The Fugene mixture was added to the cells in a dropwise manner and incubated at 37°C for 5 hours, after which the media was replaced with DMEM containing 10% FCS.

2.2.2 shRNA Lentivirus Production and Infection of Target Cells

Stable gene knockdown was achieved by infecting target cells with lentivirus containing shRNA constructs expressed from the pGIPZ plasmid (Open Biosystems, Huntsville, USA). Lentivirus was produced in 293T cells seeded at 25% confluence, with one 10cm plate per plasmid. The cells were transfected with 1µg each of p8.91 (group specific antigen and polymerase; gag-pol), pMDG.2 (Vesicular stomatitis virus-G envelope; VSV-G) and pGIPZ (containing the shRNA sequences detailed in Table 2.2) using the Fugene protocol described above. Transfection efficiency was assessed by monitoring

GFP levels 24 hours following transfection. Lentivirus was harvested 48 hours post transfection by collecting the media and passing it through a 0.45 μm filter. Aliquots of lentivirus were stored at -80°C.

Lentiviral infections were performed by adding an appropriate volume of lentivirus containing supernatant to target cells (seeded at $1x10^5$ cells per well in a 6 well plate), incubating at 37°C for 5 hours before replacing with fresh full media. Infection efficiency was assessed by monitoring the percentage of GFP positive cells 48 hours post infection with the Axiovert 100 fluorescent microscope (Carl Zeiss, Oberkochen, Germany). GFP positive cells were isolated by cell sorting with a modular flow system (Dako, Glostrup, Denmark).

Table 2.2: shRNA sequences

Target Gene	shRNA sequence (Sense, loop, antisense)
EZH2	5'-CGGAAAGAACGGAAATCTTAAATAGTGAAGCCAC AGATGTATTTAAGATTTCCGTTCTTTCCA-3'
SUZ12	5'-CGCAGTTATGCTCTTAATGCTTTAGTGAAGCCACA GATGTAAAGCATTAAGAGCATAACTGCT-3'
Non-silencing Control	5'-CTCTCGCTTGGGCGAGAGTAAGTACATCTGTGGCTTC ACTACTTACTCTCGCCCAAGCGAGAT-3'

2.2.3 Retrovirus Production and Infection of Target Cells

Retroviruses containing H-Ras^{V12} in the pWZL plasmid were produced by transfection of 293T cells using polyethylenimine (PEI; Sigma). 293T were seeded at 25% confluence, with one 10cm dish per plasmid. 4.8µg of pMLV (gag-pol), 1.6µg of pHIT (VSV-G), 8µg of pWZL (hygromycin resistance and cassette of interest) and 64µl of PEI (1mg/ml; Sigma) were added to 800µl serum free Optimem in a 1.5ml Eppendorf tube. The mixture was incubated at room temperature (22°C) for 30 minutes. During this time, the 293T cells were washed with PBS and media replaced with serum free Optimem. The transfection mixture was added to one 10cm dish of 293T cells in a dropwise manner. After 5 hours, the media was replaced with DMEM supplemented with 10% FCS. Transfection efficiency was monitored by parallel transfection with pMSCV (YFP) vector. Virus was harvested 48 hours post transfection by collecting the supernatant and passing it through a 0.45µm filter.

Target cells were seeded at 6x10^5 cells per 10cm dish the day before infection. 0.5ml of retroviral supernatant was added to each 10cm dish, incubated at 37°C for 5 hours and replaced with fresh full media. Infection efficiency was assessed by monitoring the percentage of YFP positive cells 24 hours after infection with pMSCV (YFP) retrovirus. Selection of infected cells with hygromycin (100µg/ml) was commenced 48 hours following infection and continued until no viable cells were remaining in a control population lacking the resistance cassette.

2.3 Molecular Biology Techniques

2.3.1 Isolation and Preparation of Nucleic Acids

2.3.1.1 Genomic DNA Extraction

Genomic DNA was extracted with phenol-chloroform as described previously (Miller et al., 1988). Typical cell pellets (~$1x10^6$ cells) were resuspended in 300µl nuclei lysis solution (0.1M Tris pH 8, 2mM EDTA pH 8, 0.1M NaCl), 0.33mg/ml proteinase K (Qiagen, Venlo, Netherlands) and 0.66% SDS and incubated at 60°C for 90 minutes, or until homogenous.

Protein was precipitated by adding 100µl of saturated (5M) salt solution to each sample before shaking vigorously for 2 minutes and centrifuging at 13000rpm for 15 minutes at 4°C. The supernatant was aspirated into a fresh tube and an equal volume of phenol/chloroform/isoamyl alcohol pH 8.0 25:24:1 (Sigma) was added before shaking gently for 10 minutes and centrifuging for 10 minutes at 13000rpm. The upper phase was removed to a fresh tube and an equal volume chloroform/isoamyl alcohol 24:1 (Fluka) was added. The samples were again shaken for 10 minutes and centrifuged for 10 minutes at 13000rpm before the upper phase was transferred to a fresh tube and 0.1 volumes 2M sodium acetate pH 5.6 and 1ml of ice-cold 100% ethanol was added. The solution was inverted several times to precipitate the DNA and placed at -20°C for 30 minutes. Samples were centrifuged at 13000rpm for 15 minutes to pellet the DNA before the supernatant was removed and the pellet washed in 1ml 70% ethanol. The samples were again centrifuged for 5 minutes at 13000 rpm, any

remaining alcohol was removed and the DNA pellet allowed to air dry until clear before being resuspended in 50µl 10mM Tris buffer.

The quantity and purity of the DNA was determined by measuring absorbance at 260nm (A260) and 280nm (A280) with a NanoDrop spectrophotometer (NanoDrop, Wilmington, US). An A260/A280 ratio of 1.8-2.0, indicating DNA free of contaminating protein or phenol, was obtained for all samples. Genomic DNA was stored at 4°C.

2.3.1.2 RNA Extraction

Total RNA was extracted using the RNeasy kit (Qiagen) according to the manufacturer's instructions. Cell pellets (typically ~$1x10^6$ cells) were resuspended in 350µl lysis buffer RLT and disrupted by vortexing vigorously for 15 seconds. 350µl of 70% ethanol was added before each sample was transferred to an RNeasy spin column. Following wash steps, the RNA was eluted in 30µl DEPC treated H_2O (Ambion, Austin, US). The amount and purity of the RNA was determined by measuring absorbance at 260nm (A260) and 280nm (A280) with a spectrophotometer (NanoDrop). An A260/A280 ratio of 1.8-2.0, indicating RNA free of contaminating protein or phenol, was obtained for all samples. RNA was stored at -80°C.

2.3.1.3 cDNA Synthesis

RNA samples were first treated with DNaseI (Invitrogen) to remove any contaminating genomic DNA. 1U of DNaseI (Invitrogen) was added to 300ng of RNA, 1µl of 10X DNaseI reaction buffer (Invitrogen) and made up to 10µl with DEPC-treated H_2O (Ambion). The samples were incubated for 15 minutes at room temperature (22°C) before inactivating the DNaseI by the addition of 1µl 25mM EDTA (Invitrogen) followed by a 10 minute incubation at 65°C.

1µl (200ng) of random hexamer (ABGene, Portsmouth, USA) was added before the samples were incubated at 70°C for 10 minutes and placed immediately on ice. 4µl of 5X first strand buffer (Invitrogen), 2µl of 0.1M DTT (Invitrogen), 1µl 10µM of dNTP mix (Yorkshire Biosciences, York, UK) and 1µl of SuperScript II reverse transcriptase (Invitrogen) were added to each sample. Reverse transcriptase negative controls were also made up without the addition of SuperScript II. The reaction mixtures were incubated at 25°C for 10 minutes, 42°C for 50 minutes and 70°C for 15 minutes. The final cDNA was diluted 1:5 with H_2O to a final volume of 100µl and stored at -20°C.

2.3.1.4 Bisulphite Conversion of DNA

1µg of genomic DNA was resuspended in H_2O to a final volume of 10µl before 1µl of 3M NaOH was added to a final concentration of 0.3M. The samples were incubated at 42°C for 20 minutes and cooled on ice. Sodium bisulphite solution was prepared by adding 4.56g sodium bisulphite (Sigma) to 8ml 0.3M NaOH and heating at 55°C in the dark until dissolved. Hydroquinone solution was

prepared by adding 55mg hydroquinone (Sigma) to 50ml H_2O and heating at 55°C in the dark until dissolved. 104µl of sodium bisulphite solution and 6µl of 10mM hydroquinone solution were added to each sample, followed by a 5hr incubation at 55°C in the dark.

Following bisulphite conversion, DNA was recovered using Microcon centrifugal filters (Millipore, Billerica, USA). The samples were washed three times with H_2O and desulphonated with 0.1M NaOH for 5 minutes. Following a further wash, DNA was eluted in 50µl EB (Qiagen) and stored at -20°C.

2.3.1.5 *In Vitro DNA Methylation*

100µg peripheral blood leukocyte DNA (Promega) was incubated with 1U/µg M.SssI DNA methylase (New England Biosciences, Ipswich, USA) and 0.16mM S-adeno-methionine (SAM; NewEngland Biosciences) overnight at 37°C. The following day the reaction mixture was supplemented with 0.2mM SAM and 0.065 U/µl M.SssI followed by another overnight incubation at 37°C. The fully methylated DNA was stored at 4°C.

2.3.1.6 Whole Genome Amplification

Fully unmethylated DNA was obtained by whole genome amplification of MSC 5 DNA using the Illustra GenomiPhi V2 kit (GE Healthcare, Chalfont St. Giles, UK) according to the manufacturer's instructions. 9µl of sample buffer was added to 1µl (10ng) template, heated to 95°C for 3 minutes and cooled on ice.

9µl reaction buffer and 1µl enzyme was added, and the reaction mix incubated at 30°C for 1.5 hours before inactivating the enzyme at 65°C for 10 minutes. The amplified samples, containing ~4µg DNA, were stored at 4°C.

2.3.2 Polymerase chain reaction (PCR) Based Techniques

2.3.2.1 MethyLight Quantitative Polymerase Chain Reaction (qPCR)

MethyLight reactions were performed as previously described (Weisenberger et al., 2005). qPCR reactions were conducted in a 30µl reaction volume containing 60ng of template, 0.3µM forward and reverse primers, 0.1µM probe and the remaining volume made up with TaqMan master mix No AmpErase (Applied Biosystems, Carlsbad, USA). Reactions were performed on an Eppendorf Realplex 4 (Eppendorf, Hamburg, Germany) with the cycling conditions detailed below. Primers (Operon) and probes (Applied Biosystems for MGB labelled probes and Operon for BHQ-1 labelled probes) are detailed in Table 2.3.

<u>MethyLight PCR Cycling Conditions</u>

95°C for 15 minutes

50 cycles of:

95°C for 15s

60°C for 1min

Table 2.3: Primer and probe sequences for MethyLight reactions

Element	Forward Primer	Reverse Primer	Probe
Alu-C4	5’-GGTTAGGTATAGTGGTTTA TATTTGTAATTTTAGTA-3’	5’-ATTAACTAAACTAATCTTAAAC TCCTAACCTCA-3’	5’-6FAM-CCTACCTTAACC TCCC-3’ MGB
Alu-M2	5’-GCGCGGTGGTTTACGTTT-3’	5’-AACCGAACTAATCTCGAA CTCCTAAC-3’	5’-6FAM-AAATAATCCGCC CGCCTCGACCT-3’ BHQ-1
Sat2-M1	5’-TCGAATGGAATTAATATTT AACGGAAAA-3’	5’-CCATTCGAATCCATTCGA TAATTCT-3’	5’-6FAM-CGATTCCATTCGATAAT TCCGTTT-3’ MGB
LINE1-M1	5’-GGACGTATTTGGAAAA TCGGG-3’	5’-AATCTCGCGATACGCCGTT-3’	5’-6FAM-TCGAATATTGCGTTTT CGGATCGGTTT-3’ BHQ-1

All samples were run in triplicate for each element assayed. Fold change in methylation relative to the fully methylated reference was quantified by the comparative C_t (fluorescence threshold value) method using the Alu-C4 reaction as a loading control, according to the following formulae:

$$\Delta C_{t\,(Sample)} = C_{t\,(Sample)} - C_{t\,(AluC4)}$$

$$\Delta\Delta C_t = 2\ \hat{}\ (\ \Delta C_{t\,(Meth.\ Reference)} - \Delta C_{t\,(Sample)}\ )$$

Negative control reactions containing ddH_2O and unmethylated reference DNA were performed with every reaction.

2.3.2.2 PCR for Bisulphite Sequencing

Primers were designed using the Primer3 software (http://frodo.wi.mit.edu/) against the same genomic regions assayed by MethyLight.

PCR of the Alu, Sat2 and LINE1 repetitive elements was performed in a 30µl reaction volume containing 0.3µM forward and reverse primers (Operon, Huntsville, USA), 3.5mM MgCl2, 200µM dNTPs (Yorkshire Bioscience), 1U HotStartTaq (Qiagen) and 60ng of template DNA. PCR cycling conditions and primers (Operon) were are detailed below (Table 2.4).

LINE1 PCR Cycling Conditions

95°C for 15 minutes

35 cycles of:

95°C for 30s

60°C for 20s

72°C for 30s

72°C for 5 min

Sat2 PCR Cycling Conditions

95°C for 15 minutes

35 cycles of:

95°C for 30s

55°C for 30s

72°C for 30s

72°C for 5 minutes

Table 2.4: Primer sequences for amplification of Alu, Sat2 and LINE1 repetitive elements from bisulphite converted DNA

Element	Forward Primer	Reverse Primer
Alu	5'-GGTTAGGTATAGTGGT TTATATTTGTAATTTTAGTA-3'	5'-ATTAACTAAACTAATCT TAAACTCCTAACCTCA-3'
Sat2	5'-TTTAATGGTYGYGAATGG-3'	5'-CRAATCCATAAATTATTCCA-3'
LINE1	5'-YGAGTTAAAGAAAGGGGTGA-3'	5'-CCCTCCTAACCAAATACAAA-3'

Amplification of the correct PCR products was verified by gel electrophoresis, followed by cloning and sequencing of 12-20 individual products per sample as described below.

2.3.2.3 Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR)

Total RNA was extracted and used for cDNA synthesis as above, with real-time PCR performed on an Eppendorf Realplex 4 (Eppendorf). Reactions were performed in a 25µl volume containing 0.3µM forward and reverse primers, and the remaining volume made up with 1x SYBR Green PCR mix (Applied Biosystems). PCR conditions and primers (Operon) were are detailed below (Table 2.5).

qRT-PCR Cycling Conditions

50°C for 2 minutes

95°C for 2 minutes

45 cycles of:

95°C for 15s

60°C for 1 min

Table 2.5: Primer sequences for qRT-PCR reactions

Gene	Forward primer	Reverse primer
DNMT1	5'-CTGTACCGAGTTGGTGATGGTGTG-3'	5'-CTCTGGGTACAGGTCCTCATCCAC-3'
EZH2	5'-TTTTTGCCAAGAGAGCCATC-3'	5'-AAGGCAGCTGTTTCAGAGGA-3'
SUZ12	5'-CCGTCCACAAGAAATGGAAG-3'	5'-TCATGACATGGAGATTCCAGA-3'
EED	5'-CCAGACGGACACTCTGGTG-3'	5'-GAATGATCCATACCACAGGACA-3'
GAPDH	5'-GGAGTCAACGGATTTGGTCGTA-3'	5'-GGCAACAATATCCACTTTACCAGAGT-3'
DKK2	5'-ATGTGCTGCCCCAGTACC-3'	5'-GTGCCGAGTACCATCCAGAG-3'
EN1	5'-CTGGGTGTACTGCACACGTT-3'	5'-GCTTGTCCTCCTTCTCGTTC-3'
HOXB3	5'-CTGGGGCTCGATGTGAATA-3'	5'-CTCCTCCGGGGTCTGTTC-3'
TBX2	5'-GCTGAAGCTGACCAACAACA-3'	5'-GCTGGTACTTGTGCATGGAG-3'
RBP4	5'-ACCCTGCCAAGTTCAAGATG-3'	5'-TCTGTGTCGACGATCCAGTG-3'

Relative expression levels were quantified by the comparative C_t (fluorescence threshold value) method using GAPDH as a loading control, according to the following formulae:

$$\Delta C_{t\,(Sample)} = C_{t\,(Sample)} - C_{t\,(GAPDH)}$$

$$\Delta\Delta C_t = 2\ \hat{}\ (\ \Delta C_{t\,(Reference)} - \Delta C_{t\,(Sample)}\ )$$

All samples were run in triplicate for each gene assayed alongside negative control reactions using ddH_2O and reverse transcriptase negative cDNA samples.

2.3.3 PCR Product Purification and Cloning

2.3.3.1 Agarose Gel Electrophoresis

PCR products and restriction enzyme digests were visualized by agarose gel electrophoresis. Typically, 1.5% weight / volume (w/v) agarose (Sigma) gels were made up in 0.4M tris-acetate, 0.01M ethylene diamino tetraacetic acid (TAE; National Diagnostics, Atlanta, USA) containing 0.04 mg/ml ethidium bromide (Sigma). DNA samples were loaded using 6x DNA loading buffer (Fermentas, Burlington, Canada) and run in TAE at 100V. Bands were visualized using Sygene G:Box UV transilluminator (Syngene, Frederick, USA) and run alongside a 1kb DNA ladder (Fermentas) to estimate the size of the DNA molecules.

2.3.3.2 Agarose Gel Extraction of DNA

Appropriate bands were excised from the gel with a clean scalpel and placed in a 1.5ml Eppendorf tube. DNA was extracted using the QIAquick Gel Extraction Kit (Qiagen) as described by the manufacturer. After weighing the gel fragment, an appropriate volume of buffer QG was added (3µl for ever µg of gel) and the solution was incubated at 50°C for 10 minutes to dissolve the gel. Isopropanol was added to the solution (1µl for every µg of gel) and the samples were loaded onto the QIAquick spin column and washed as described by the manufacturer. The DNA was eluted in 20µl ddH_2O and used immediate in a ligation reaction or stored at -20°C

2.3.3.3 Ligation of PCR Products into pGEM-T

3µl of gel extracted PCR product was added to 1µl (50ng) of pGEM-T Easy vector (Promega), 1µl of (3 Weiss units) T4 DNA ligase (Promega), and 5µl of 2X reaction buffer (Promega). The ligation reaction was allowed to proceed at room temperature (22°C) for 1 hour, and used immediately for bacterial transformation or stored at -20°C. A negative control (containing H_2O instead of insert) was always conducted in parallel to monitor self-ligation of the vector.

2.3.3.4 Other Ligations

Ligations were performed in a 20µl volume containing 1µl (20 units) of T4 DNA ligase (New England Biosciences), 2µl of 10x DNA ligase buffer (New England

Biosciences) and 50 to 200ng of vector. The amount of insert added to the ligation mixture (A_{Insert}) was calculated using the following formula, with a ratio of 8:1 (insert:vector) optimal.

$$A_{Insert} = [(A_{Vector} * L_{Insert}) / L_{Vector}] * Ratio$$

Where L_{Insert} and L_{Vector} denote the length of insert and vector respectively (measured in kilobases), A_{Insert} and A_{Vector} are the amount of insert and vector respectively (measured in nanograms) and 'Ratio' is an integer value of the insert:vector ratio (for example, 8).

Ligation reactions were conducted at room temperature (22°C) for 1 hour or at 4°C overnight for optimal efficiency. Vector controls containing no insert were run in parallel to monitor the incidence of vector self ligation.

2.3.4 Plasmid Preparation and Purification

2.3.4.1 Bacterial Transformation

5µl of ligated plasmid was added to One Shot TOP10 chemically competent *E. coli* cells (Invitrogen), mixed gently and incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 30 seconds and rapidly cooled on ice. 250µl of pre-warmed SOC medium (Invitrogen) was added to the cells and the vials were shaken at 37°C for 1hr at 225rpm. 80µl of the culture was spread onto LB agar plates (Invitrogen) supplemented with 100µg/ml ampicillin (Sigma) to select for successfully transformed bacteria.

2.3.4.2 Minipreps

A single bacterial colony was picked and grown overnight at 37°C in 5ml of Luria-Bertani broth (Invitrogen) supplemented with 100µg/ml ampicillin (LB^{amp}). Plasmids were extracted with a MiniPrep kit (Qiagen) according to the manufacturer's instructions. Bacteria were pelleted by centrifugation at 3500rpm for 15 minutes. The pellet was resuspended in 250µl buffer P1, lysed and neutralised through an alkaline lysis method according to the protocol. The resultant solution was centrifuged at 13000rpm for 10 minutes and the supernatant was transferred to a Qiagen column. Wash steps were performed according to the manufacturer's protocol and the plasmid DNA was eluted in 50µl of ddH_2O. Typical minipreps yielded approximately 10µg DNA.

2.3.4.3 Maxipreps

When larger amount of plasmid were required for cell transfections and virus production, purification was conducted using a maxiprep kit (Qiagen). 100ml of bacterial culture was grown overnight by inoculating LB^{amp} with a single colony picked from an agar plate. The following day, the bacterial suspension was pelleted by centrifugation at 3500rpm for 15 minutes at 4°C. The pellet was resuspended in 10ml buffer P1, lysed by the addition of 10ml buffer P2 at room temperature (22°C) for 5 minutes and neutralised with 10ml chilled buffer P3. The mixture was incubated on ice for 20 minutes and centrifuged at 3500rpm for 30 minutes at 4°C. The supernatant was added to a Qiagen maxiprep column that had been previously equilibrated with 10ml buffer QBT. Following wash steps as

directed by the manufacturer, the DNA was eluted and precipitated with isopropanol. The DNA was pelleted and washed as described in the protocol before being resuspended in 300µl H_2O overnight at 4°C.

2.3.4.4 Restriction Enzyme Digestion

Diagnostic restriction enzyme digestions were performed using 10 units of enzyme in a reaction volume of 25µl containing the appropriate 1x reaction buffer (New England Biosciences) and 1µg or less of DNA. Reactions were typically performed for 1-2 hours at 37°C. Successful digestion was confirmed by gel electrophoresis of the products, with appropriate bands excised and gel extracted when required.

2.3.5 DNA Sequencing

2.3.5.1 Sanger Sequencing

Sanger sequencing of PCR products cloned in the pGEM-T easy vector was performed by the Scientific Support Service (WIBR, UCL, London). Chain termination reactions were conducted using T4 sequencing primer and BigDye terminator (Applied Biosystems), with reaction products sequenced on an ABI PRISM 3700 sequencer (Applied Biosystems).

2.3.5.2 Pyrosequencing

Methylation analysis of selected promoters was performed by pyrosequencing. PCR reactions were conducted in a 30µl reaction volume containing 1.5mM $MgCl_2$ (Qiagen), 0.2mM dNTPs (Yorkshire Biosciences), 5 pmoles forward and reverse PCR primers (Operon), 0.2 units of HotStarTaq (Qiagen) and 1x PCR buffer (Qiagen). PCR cycling conditions and primers sequences (Operon) are detailed below (Table 2.6).

PCR Cycling Conditions

95°C for 15 minutes

45 cycles of:

95°C for 20 sec

50°C for 20 seconds

72°C for 20 seconds

72°C for 5 minutes

Table 2.6: Primer sequences for pyrosequencing PCR product amplification from bisulphite converted DNA

Gene	Forward Primer	Reverse Primer
DKK2	5'-GTAAAGAGGATTGGGGAGAGAGTA-3'	5'-BIO-CCCCTAACTCACAAAAAACAAC-3'
EN1	5'-TGTGGGTTTGGGTGTGTGTA-3'	5'-BIO-CTCCCACTACCAAAAAAAATCCA-3'
RBP4	5'-TAGGTAGGGTTTTTGGGGAATT-3'	5'-BIO-CAAACCCCTACTATCATCCTT-3'
TBX2	5'-GATTATAGAAGAGGGGGTTTTAATTATAGG-3'	5'-BIO-ACCCCCCCCTAAACTTTCCA-3'

10µl of the biotin labelled PCR product was made up to 40µl with ddH_2O before 2µl of straptavidin-coated sepharose beads (GE Healthcare) and 38µl of binding buffer (Qiagen) was added to each sample. The PCR product was immobilized to the beads by shaking for 5 minutes at room temperature (22°C).

The beads were captured on a Vacuum Prep Tool (Qiagen) and subjected to strand separation by treatment with 70% ethanol, denaturation solution (0.2M NaOH) and wash buffer (Qiagen). The beads were released into a PSQ HS 96 plate containing 12µl of 0.3uM sequencing primer per well. Primer annealing was achieved by heating the plates to 80°C for 2 minutes before cooling the samples to room temperature (22°C).

The pyrosequencing reaction was conducted on a PyroMark MD sequencer (Qiagen). A bisulphite conversion control was included in each assay. The methylation status of each CpG site was analyzed using the Pyro Q-CpG software (Qiagen), with average methylation values calculated for each PCR product. Sequencing primers and dispensation orders for the reactions are detailed below (Table 2.7).

Table 2.7: Sequencing primers and nucleotide dispensation orders for pyrosequencing

Gene	Sequencing Primer	Dispensation Order
DKK2	5'-GAGAGAGTAGAGAGAGAGAA-3'	GATGTCGATGTCAGTCTGTCGATGTCGTCATGTCGTATGTC GTTAGTGATCGTAGTCG
EN1	5'-AAATAGTAAAATTTGAAAGG-3'	GTCGTAGTCGTAGTATGTCAGTCGTGATCGTGATCGTGTGA TCGTTAGTCTGTC
RBP4	5'-TTTTTGGGGAATTTTGGA-3'	AGTCGTATGTCGTCAGTAGTTCTGTCGTATGATTAAGTATC GTATCGTT
TBX2	5'-GAGAAGAAATTAAAAAAGTGG-3'	ATGCTATGTCGAGAGAGATGTCGTGATTCGAGAGAGTATTC AGTTCGTGATCGATATGTCGAA

2.3.6 *Western Blot*

Cells were lysed in RIPA buffer containing protease and phosphatase inhibitor cocktails (Sigma), as described below.

Cell Lysis (RIPA) Buffer

150 mM NaCl

50mM Tris-HCl pH 7.5

1% (v/v) NP40

0.5% (w/v) Deoxycholic acid

0.1% (w/v) Sodium dodecyl sulfate (SDS)

1:100 Protease inhibitor cocktail

1:100 Phosphatase inhibitor cocktail I

1:100 Phosphatase inhibitor cocktail II

50-100µl of lysis buffer was added directly to cells in a 10cm dish that had been washed with 5ml cold PBS. The cells were scraped, aspirated into a 1.5ml Eppendorf tube and incubated for 30 minutes on ice. Samples were centrifuged at 13000rpm for 30 minutes at 4°C and the supernatant transferred to a fresh tube.

Protein was quantified by the Bradford assay. 1µl of cell lysate was added to 1ml Bradford solution (BioRad Laboratories, Hercules, USA) and 100µl H_2O. Absorbance at 595nm was measured on a Fluoroskan spectrophotometer (Thermo Fisher) and protein concentration calculated on the basis of a regression

line obtained from known protein standards comprising dilutions of bovine serum albumin (BSA) solution (Promega).

Typically, 20µg of protein was mixed with 4x sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample loading buffer, heated at 100°C for 3 minutes and allowed to cool. SDS-PAGE loading buffer composition is described below.

SDS-PAGE loading buffer

250mM Tris-HCl pH 6.8

8% (w/v) SDS

10% (v/v) Glycerol

5% (v/v) Beta-mercaptoethanol

0.05% (w/v) Bromophenol blue

Proteins were resolved on a 12% SDS-PAGE gel (comprising a stacking gel and resolving gel described below) run at 100V for two hours in SDS-PAGE running buffer (National Diagnostics) and transferred to an immobilon membrane (Millipore) using a Transblot SD semi-dry transfer unit (BioRad) and transfer buffer (National Diagnostics) at 20V 500mA for 1hr. Membranes were blocked in 5% (w/v) milk in PBS containing 0.05% (v/v) Tween 20 (Sigma) for 1hr and incubated with primary antibody overnight at 4°C. Three 10 minute washes in PBS 0.05% Tween were performed before probing the membrane with secondary horseradish peroxidase (HRP) conjugated antibody at room temperature (22°C)

for 1 hour. Following three 10 minute washes in PBS 0.05% Tween, excess wash buffer was removed from the membrane before it was covered with ECL detection reagent (GE Healthcare) and incubated at room temperature (22°C) for 1 minute. Excess ECL was drained and bands were visualized with high performance autoradiography film (GE Healthcare). Densitometric quantification of bands was performed on scanned images using the Scion Image software (Scion Corporation, Frederick, USA). Antibodies and concentrations used for western blots are described below (Table 2.8).

Stacking Gel

1.4 ml ddH_2O

0.33 ml 30% (w/v) acrylamide

0.25 ml 1M Tris pH 6.8

0.02 ml 10% (w/v) SDS

0.02 ml 10% (w/v) ammonium persulfate

0.002 ml TEMED

12% Resolving Gel

3.3 ml ddH_2O

4 ml 30% (w/v) acrylamide

2.5 ml 1.5M Tris pH 8.8

0.1 ml 10% (w/v) SDS

0.1 ml 10% (w/v) ammonium persulfate

0.004 ml TEMED

Table 2.8: Antibodies used in this study for western blots

Antigen	Antibody	Species	Monoclonal/ Polyclonal	Dilution	Source
H-Ras	259	Rat	Monoclonal	WB 1:100	Santa Cruz
Phospho-p44/42 MAP Kinase	9101	Rabbit	Monoclonal	WB 1:1000	Cell Signaling
Total ERK1/2	06-182	Rabbit	Monoclonal	WB 1:5000	Upstate
GAPDH	RGM2	Mouse	Monoclonal	WB 1:10000	Advanced Immunochemicals
Anti-Rat HRP	P0450	Rabbit	Polyclonal	WB 1:5000	Dako
Anti-Mouse HRP	P0447	Goat	Polyclonal	WB 1:5000	Dako
Anti-Rabbit HRP	P0448	Goat	Polyclonal	WB 1:5000	Dako
EZH2	3147	Mouse	Monoclonal	WB 1:1000	Cell Signaling
EED	SC-28701	Rabbit	Polyclonal	WB 1:1000	Santa Cruz
SUZ12	3737	Rabbit	Monoclonal	WB 1:1000	Cell Signaling
H3K27me3	07-449	Rabbit	Polyclonal	WB 1:2000	Upstate
Histone H3	9715	Rabbit	Polyclonal	WB 1:1000	Cell Signaling

2.3.7 Chromatin Immunoprecipitation (ChIP)

ChIP was performed using the Histone ChIP Kit (Diagenode, Liege, Belgium) according to the manufacturer's instructions. Each 10cm dish of cells was washed in PBS and crosslinked by the addition of 8ml serum free DMEM containing 1% formaldehyde (Sigma) for 10 minutes at room temperature (22°C). The crosslinking reaction was neutralized by the addition of 800µl 1.25M glycine and the plates were washed twice with ice cold PBS before cells were scraped, aspirated to a 1.5ml Eppendorf tube and centrifuged at 2000rpm for 5 minutes. The supernatant was removed and the cell pellet resuspended in an appropriate volume of buffer A, to a final concentration of $1.25x10^5$ cells per 10µl volume.

Chromatin was sheared in a Bioruptor sonicator waterbath (Diagenode) for 30 minutes at high power, with a cycle interval of 30 seconds on and 30 seconds off. Samples were centrifuged at 13000rpm for 5 minutes to pellet debris, with the supernatant containing the sheared chromatin retained. Successful shearing of this chromatin to an average size of 200-600bp was confirmed by running 5µl of the sample on a 1% agarose gel. Unused chromatin was snap frozen in liquid nitrogen and stored at -80°C.

Each IP was performed using 10µl of sheared chromatin, corresponding to $1.25x10^5$ cells, with 10µl of each sample stored at -20°C to use as an input control. The samples were diluted in IP incubation mix and pre-cleared with 20µl of the pre-blocked beads for 1 hr at 4°C. The beads were pelleted by centrifugation at 3000rpm for 2 minutes and the supernatant retained. 1.5µg of

antibody was added for each IP and incubated overnight at 4°C on a rotating wheel. IP was conducted using antibodies against H3K27me3 (Upstate, Millipore), total H3 (Abcam, Cambridge, UK) and an IgG isotype control (Becton Dickinson, Oxford, UK).

The following day, 20µl of pre-blocked beads were added to each sample before incubating at 4°C for 1 hour on a rotating wheel. The beads were pelleted by centrifugation and washed as directed in the protocol. The precipitated chromatin was eluted in 400µl buffer C, with input samples processed alongside precipitated samples from this point onwards. Crosslinking was reversed by adding 16µl 5M NaCl to each sample and incubating at 65°C for 4 hours with rotation. Finally, DNA was phenol/chloroform purified, precipitated in ethanol overnight, washed and resuspended in 50µl H_2O as described in the protocol.

2.3.7.1 ChIP-PCR

5µl of IP DNA and 100ng of input DNA was used as template for each qPCR reaction conducted in a volume of 25µl containing 0.3µM forward and reverse primers, and 1x SYBR Green PCR mix (Applied Biosystems). PCR Cycling conditions and primer sequences (Operon) are described below (Table 2.9).

PCR Cycling Conditions

50°C for 2 minutes

95°C for 2 minutes

45 cycles of:

95°C for 15s

60°C for 1 min

Table 2.9: Primer sequences for ChIP-PCR reactions

Gene	Forward primer	Reverse primer
DKK2	5'-AGCTAGATCCCCCTTTTCCA-3'	5'-GTGCTGCTTCTGCTTCTGC-3'
EN1	5'-CCCTCCCTCCACTACACACA-3'	5'-TTGAAGGATGGGTGGTTGTT-3'
HOXB3	5'-GCCTCAGGGTCTCTGGTCTA-3'	5'-CCGAGCTTTTTGGAGATGTG-3'
DCN	5'-CTGGTGGACAGGGAGAAAGA-3'	5'-TGTCCTGAAAGAAAGTCATGGTT-3'

All samples were run in triplicate alongside negative control reactions containing ddH_2O. Percentage input for the H3K27me3 and total H3 IPs were calculated as:

$$\% \text{ input} = 2 \text{ ^ } (C_{t\,(Input)} - C_{t\,(ChIP)}) * 100$$

Where $C_{t\,(Input)}$ and $C_{t\,(ChIP)}$ are threshold values for the input and IP reactions respectively.

2.3.8 *Flow Cytometry*

2.3.8.1 Fluorescence Assisted Cell Sorting (FACS) for Methylated DNA

Cells were harvested and fixed in 70% ethanol at -20°C for 30 minutes, pelleted by centrifugation at 2500rpm for 5 minutes and washed in 1ml PBS (Sigma). Samples were permeabilised with 500µl 0.2% Triton-X-100 (Sigma) for 15

minutes at room temperature (22°C). Histones were removed by treating the cells with 500µl 2M HCl at 37°C for 25 minutes, before neutralising the acid with 1ml 0.1M Tris-HCl pH 8.5 for 5 minutes at room temperature (22°C). Although this treatment is likely to hydrolyse RNA, the cells were further treated with 200µl 0.1mg/ml RNase A (Sigma) for 30 min at 37°C. Following a wash with 1ml PBS supplemented with 1% BSA (Sigma) and 0.1% Tween 20 (Sigma) (PBST-BSA), cells were blocked in 1ml PBST-BSA for 20 minutes at 37°C. Samples were incubated with mouse anti-5-methylcytosine antibody (Eurogentec, Seraing, Belgium) for 1 hour at 37°C in a volume of 100µl. The cells were washed three times in PBS before incubation for 45 minutes at 37°C with anti-mouse secondary antibodies conjugated to fluorescein isothiocyanate (Vector Labs, Burlingame, USA) at a concentration of 7.5µg/ml in a volume of 50µl. Samples were washed three times in PBS and resuspended in 500µl 50µg/ml propidium iodide (Sigma) for 45 minutes at 37°C. Cell suspensions were analysed on a FACScalibur cell sorter (Becton Dickinson, Oxford, UK) using CellQuest analysis software. Full details of antibodies used for FACS are described below (Table 2.10).

Table 2.10: Antibodies used for FACS analysis of nuclear methylcytosine content

Antigen	Antibody	Species	Monoclonal/ Polyclonal	Dilution	Source
IgG1 control	555746	Mouse	Polyclonal	FACS 1:1000	BD
5'-Methyl cytidine	33D3	Mouse	Monoclonal	FACS 1:1000	Eurogentec
Anti-Mouse FITC		Goat	Polyclonal	FACS 1:200	Vector Labs

2.3.8.2 FACS for Cell Cycle Analysis

Cells were harvested and fixed in 70% ethanol at -20°C for 30 minutes, pelleted by centrifugation at 2500rpm for 5min and washed in 1ml PBS (Sigma). Samples were permeabilised with 500µl 0.2% Triton-X-100 (Sigma) for 15 minutes at room temperature (22°C) before being washed in PBS and resuspended in 500µl 50µg/ml propidium iodide (Sigma) and 0.1mg/ml RNase A (Sigma) for 30 minutes at 37°C. Cell suspensions were analysed on a FACScalibur (Becton Dickinson, Oxford, UK) using CellQuest analysis software. Percentage of tetraploid cells in each population was calculated using the ModFit software (Verity Software House, Topsham, USA).

2.3.9 Karyotypic Analysis

Fully karyotyping of the MSC 5 cell line was conducted by Helena Kempski and Steve Chatters (Institute of Child Health, Great Ormond Street Hospital, London) as described previously (Kempski et al., 2006).

2.4 Bioinformatics

2.4.1 Gene Expression Microarray (GEM) Data Analysis

Gene expression microarray data for the 5 hit stem cell model was obtained previously (Funes et al., 2007). cDNA from each cell line (with the exception of MSC 2E6) was hybridized to Affymetrix hg-u133+2 GeneChips (Affymetrix, Santa Clara, USA) in triplicate and processed to obtain log expression values (Funes et al., 2007). Significant changes in gene expression between MSC 0 and MSC 5 were determined by FDR corrected T-Test of log expression values between conditions, with $q \leq 0.01$ considered significant. In order to adjust gene expression data for cell proliferation rate, the statistical analysis described above was also repeated on log expression values for all genes normalized against expression of the proliferation-linked marker PCNA.

In order to focus the analysis on relevant genes involved in epigenetic regulation, a list of epigenetic regulators was compiled from an extensive literature review, comprising 967 probe sets for 341 genes. The list included DNA methyltransferases, histone deacetylases, histone acetyltransferases, histone methyltransferases, histone demethylases, methyl-cytosine binding proteins, ATP dependent chromatin remodelling proteins, imprinted genes, micro-RNA processing genes, nucleosome assembly proteins, polycomb proteins and cytosine deaminases (Table A2). GEM data for this list was extracted from the total array data for further analysis.

2.4.2 Gene Expression Heat Maps

Heat maps were created with the dChip microarray analysis package (http://www.biostat.harvard.edu), using the relevant probe sets and log expression values. These were used to visually represent differential changes in gene expression, and display expression of each probe set in units of standard deviation across all replicates. The colour corresponds to the direction of change (down regulation: blue, up-regulation: red), with intensity correlating with magnitude of the change (scale in units of standard deviation is displayed under each figure).

2.4.3 Gene Set Enrichment Analysis (GSEA)

GSEA is an unbiased computational method used to determine whether a particular set of genes are significantly up- or down-regulated in comparison to a ranked list of all genes in a GEM experiment (Subramanian et al., 2005). The enrichment score measures the skew of the gene set within the order of genes sorted by expression difference between two samples. Significance is determined by the false discovery rate (q value), a measure of enrichment scores greater than 1000 enrichment scores generated from a random gene sets of equal size to the gene set of interest.

2.4.4 Statistical Analyses

Student's T-tests were conducted using Microsoft Excel as two tailed, unequal sample variance, with $p \leq 0.05$ considered statistically significant. FDR correction for multiple testing and Mann-Whitney U tests were conducted using the statistical package R. Error bars are standard error of the mean (SE), calculated as standard deviation divided by the square root of the number of samples.

Chapter 3: Transformation of mesenchymal stem cells can induce repetitive element hypomethylation

3.1 Background

Many cancer types show hallmark epigenetic changes including characteristic histone modifications, gene specific hypermethylation and global hypomethylation (Feinberg and Tycko, 2004). Of these, the first to be described was genome wide hypomethylation, a decrease in the overall genomic methylcytosine content from approximately 4% in normal tissues to 2-3% in cancers. This change was first observed in a 1983 study comparing lung and colon carcinomas to adjacent normal tissue, demonstrating that overall genomic methylcytosine levels were lower in cancer tissues (Feinberg and Vogelstein, 1983). Since this pioneering study, the observation has been reproducibly repeated in comparisons of a wide range of cancers and normal tissues using a variety of different techniques (Ehrlich et al., 2002; Feinberg et al., 1988; Gama-Sosa et al., 1983a; Narayan et al., 1998; Qu et al., 1999; Weisenberger et al., 2005).

The range of methods available for the study of global methylation has also expanded considerably since initial studies in the field. Traditionally, global methylcytosine content was measured using high performance liquid chromatography (HPLC) of DNA digested to individual deoxynucleosides (Gama-Sosa et al., 1983a; Gama-Sosa et al., 1983b; Kuo et al., 1980). Although

this method gives an absolute quantitative measure of methylcytosine content, it is time consuming and requires a large amount of DNA. Several alternative methods, based on measuring DNA methylation in repetitive elements as a proxy for global methylation, have since been developed. These include:

1. Digestion of genomic DNA with methylation sensitive restriction enzymes followed by Southern blot analysis against repetitive elements, with differences in digestion pattern indicative of changes to DNA methylation (Jurgens et al., 1996; Narayan et al., 1998). Although reproducible and easier than HPLC, this method does not give a quantitative measure of genomic methylcytosine content and also requires a large amount of DNA.
2. Bisulphite sequencing, based on bisulphite conversion of unmethylated cytosine residues to uracil, which also gives an absolute readout of DNA methylation at individual CpG dinucleotides (Frommer et al., 1992; Shapiro et al., 1973). However, this technique is only practical in assessing methylation at unique loci, not on a genome wide scale unless a repetitive element is assayed. Additionally, the technique is very time consuming and expensive due to the number of PCR products that have to be individually cloned and sequenced.
3. A methylation sensitive PCR (MSP) and the quantitative PCR (qPCR) based method MethyLight (Trinh et al., 2001; Weisenberger et al., 2005). This method utilizes primers which selectively amplify methylated or unmethylated sequences in bisulphite converted DNA. For MethyLight, methylation levels are expressed in relation to a fully methylated or fully

unmethylated control. These methods require a very small amount of starting DNA and provide high throughput, quantitative results.

4. Monoclonal antibodies against 5-methylcytosine have been developed and used to reliably stain methylated DNA *in vitro* (Bensaada et al., 1998; Habib et al., 1999). As well as being applied to immunohistochemistry, such antibodies may be used for more quantitative measurement of methylated DNA content in a cell by fluorescence assisted flow cytometry (FACS) where individual cells can be stained and assayed. However, this is a relatively new method, the accuracy of which has not yet been fully established.
5. Methylated DNA immunoprecipitation (MeDIP). Methylated regions of the genome can be precipitated from whole genomic DNA using antibodies against methyl-cytosine, with the identity of the precipitated DNA confirmed by microarray analysis or high throughput sequencing (Weber et al., 2005). As well as providing a measure of global methylation levels, these methods allow regions from which methylation is lost to be identified. However, due to the current high cost of these techniques, they have not been routinely used to address global hypomethylation in cancer.

Despite these numerous observations of genome-wide hypomethylation in cancer, relatively little is known about the causes of this loss of methylation. Most studies to date have been conducted in primary cancer tissues compared only to corresponding matched or unmatched normal tissues. In doing so, a specific link between a particular genetic or epigenetic change and the induction

of global hypomethylation in these cancers can not be established due to variation in the cellular composition or the genetic and epigenetic background of tumours compared to corresponding normal tissues. Furthermore, although matched cases and control samples remove the effect of some of this genetic variation, no convincing studies have been conducted in which hypomethylation has been induced in a specific cellular population through transformation with defined genetic or epigenetic changes.

The step-wise model of MSC transformation developed in our laboratory provides an excellent model in which to study epigenetic changes such as the induction of global hypomethylation. This model allows a direct comparison of normal cells (parental MSC) to equivalent transformed cells derived from the same lineage. Furthermore, by inducing transformation in a step-wise manner, a link between these oncogenic 'hits' and subsequent epigenetic changes can be established. I hypothesize that global DNA hypomethylation occurs during step-wise transformation in MSC, and aim to utilize a number of different methods to assess genome-wide methylation levels at each step in the model. An investigation of genome-wide methylation levels in these cells could therefore suggest a causal relationship between the genetic change and induction of hypomethylation. Furthermore, this study will also determine whether this change is an early or late event during the transformation process.

3.2 Aims

1) To test and optimise a number of methods for assaying genome-wide DNA methylation levels, namely:
    a. Antibody staining of 5-methylcytosine followed by FACS analysis of fluorescence.
    b. MethyLight, a qRT-PCR based assay of repetitive element methylation.
2) Utilize these methods to establish whether genome wide hypomethylation is observed in the MSC model and identify the step at which it occurs.

3.3 Results

3.3.1 FACS Analysis of Global Methylation

In order to obtain a direct measure of total DNA methylation in the MSC cell lines, the feasibility of a novel flow cytometry based method for measuring global methylation was explored. In this method, cells are permeabilised and incubated with a primary antibody against methylated cytosine. A fluorescein isothiocyanate (FITC) conjugated secondary antibody is used to label the primary antibody, such that green fluorescence directly correlates to the amount of methylated DNA within the nucleus. To control for total amount of DNA, co-staining of total genomic DNA with propidium iodide (PI) is also carried out in parallel. Both fluorescence readings are obtained at the single cell level with a flow cytometer.

Extensive optimisation of this method was carried out using MSC 5 cells (Figure 3.1). Briefly, 150,000 cells per condition were fixed, permeabilized and stained for methylcytosine (green) and total DNA (red), with both fluorescence measurements recorded by FACS on FL1 (methylcytosine) and FL2 (total DNA) channels. Initially, methylcytosine staining was not distinguishable from background fluorescence (Figure 3.1a). In order to overcome this problem, the protocol was modified to include a treatment with hydrochloric acid to remove histones, which resulted in a significant increase in specific staining (Figure 3.1b) (Habib et al., 1999).

Primary and secondary antibody concentrations were also optimised by titration. MSC 5 cells were incubated with increasing concentrations of primary

antibody, ranging from 0.05 - 4μg, followed by incubation with 7.5μg/ml secondary antibody. Efficiency of staining was determined by monitoring the geometric mean (GeoMean) green fluorescence (FL1) of the total cell population. These data indicated that the primary antibody saturates when present at amounts greater than 1μg (Figure 3.1c). Since optimal results are obtained when the primary antibody is not in excess, 0.1μg of antibody was used in further experiments. Titration of the secondary antibody was conducted in a similar manner. In this instance, fluorescence continued to increase with higher concentrations of antibody, indicating that the saturation point had not been reached at the highest concentration (15μg/ml) (Figure 3.1d). A secondary antibody concentration of 7.5μg/ml was chosen for further experiments.

Following optimisation of this method, analysis of genome-wide methylation levels in the MSC cell lines was conducted, comparing untransformed MSC 4 with transformed MSC 5. As a positive control, MSC 4 cells were treated with 5μM 5-aza-deoxycytidine for four days to induce hypomethylation and assayed alongside the experimental samples. FL1 signals above the levels from a matched isotype control antibody were obtained for all samples, indicating specific staining of methylcytosine in these cells (Figure 3.2c). GeoMean FL1 measurements were analysed for the total cell populations as well as for cells in G1, M and G2 phases of the cell cycle (Figure 3.2d). These data indicated that the method was able to detect a significant decrease in methylcytosine levels when MSC 4 cells were treated with the hypomethylating agent 5-aza-deoxycytidine. However, no significant difference in methylation was evident between MSC 4 and transformed MSC 5 cells. To control for the amount of total DNA in these cells, the FL1:FL2 ratio (methyl-cytosine : total

DNA) was calculated for these cell populations (Figure 3.2e). These results also indicated a significant decrease in genome-wide methylation in the 5-aza-deoxycytidine treated cells, but detected no significant difference between MSC 4 and MSC 5 cells at any stage of the cell cycle.

Figure 3.1

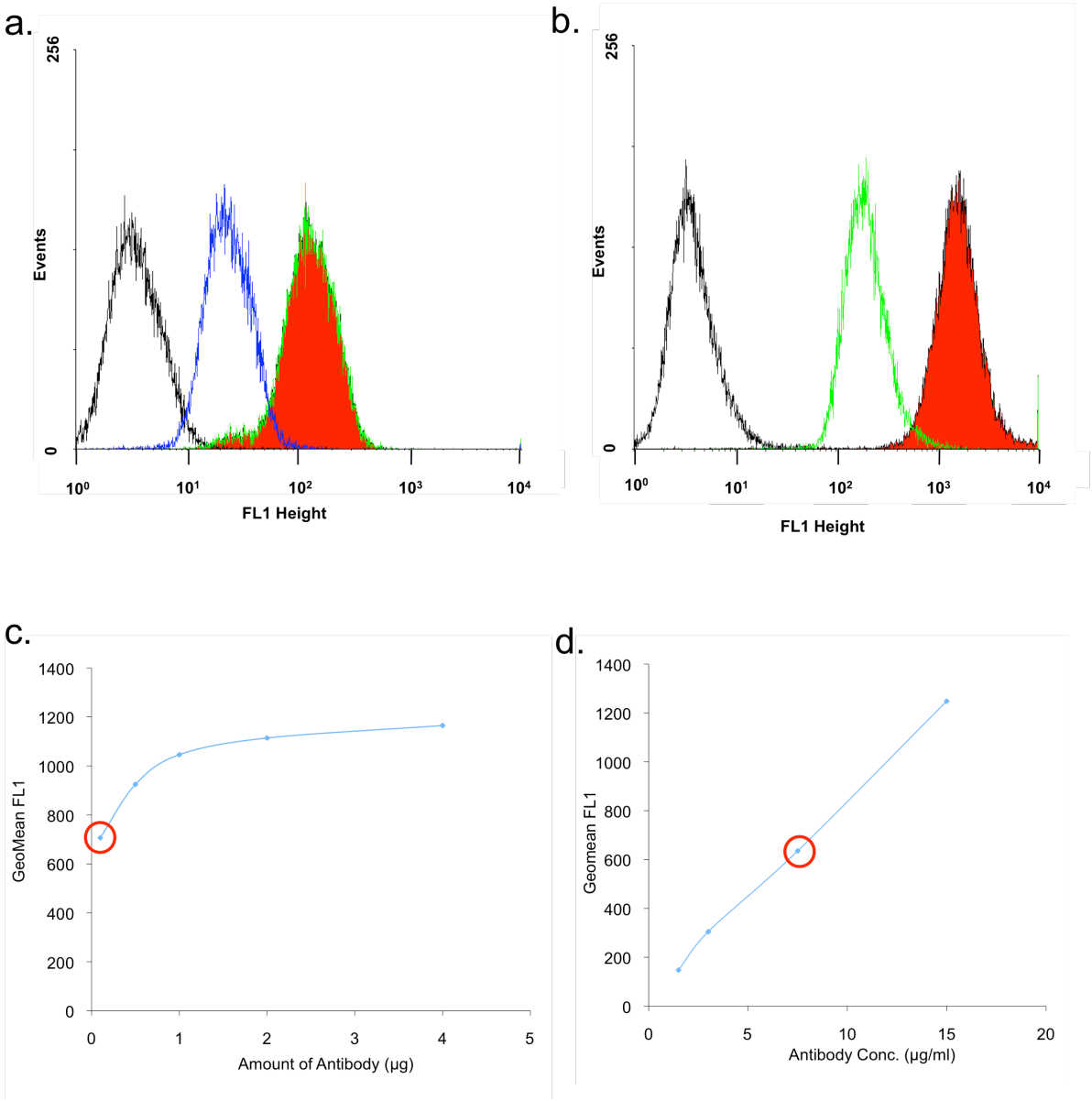


Figure 3.1: Optimisation of a FACS based method of measuring genomic 5-methylcytosine content. MSC 4 were grown as normal, fixed in ethanol and permeabilized with Triton-X. **(a)** MSC 4 stained with mouse anti-methylcytosine primary antibody followed by anti-mouse FITC secondary antibody. **(b)** MSC 4 incubated with 2M HCl for 25 minutes before the addition of primary and secondary antibody. Flow cytometry analysis shows FL1 signal (methyl-cytosine) for unstained (black), secondary antibody only (blue), isotype control (green), and double stained (red) cells. **(c)** MSC 4 stained with increasing amounts of mouse anti-methylcytosine primary antibody followed by anti-mouse FITC secondary (8 µg/ml). **(d)** MSC 4 stained with 1µg mouse anti-methylcytosine primary antibody followed by increasing concentrations of anti-mouse FITC secondary antibody. GeoMean FL1 was recorded by flow cytometry and corrected for background fluorescence as measured by staining with an equivalent amount of IgG1 isotype control. Optimal concentrations used for further experiments are indicated with red circles.

Figure 3.2

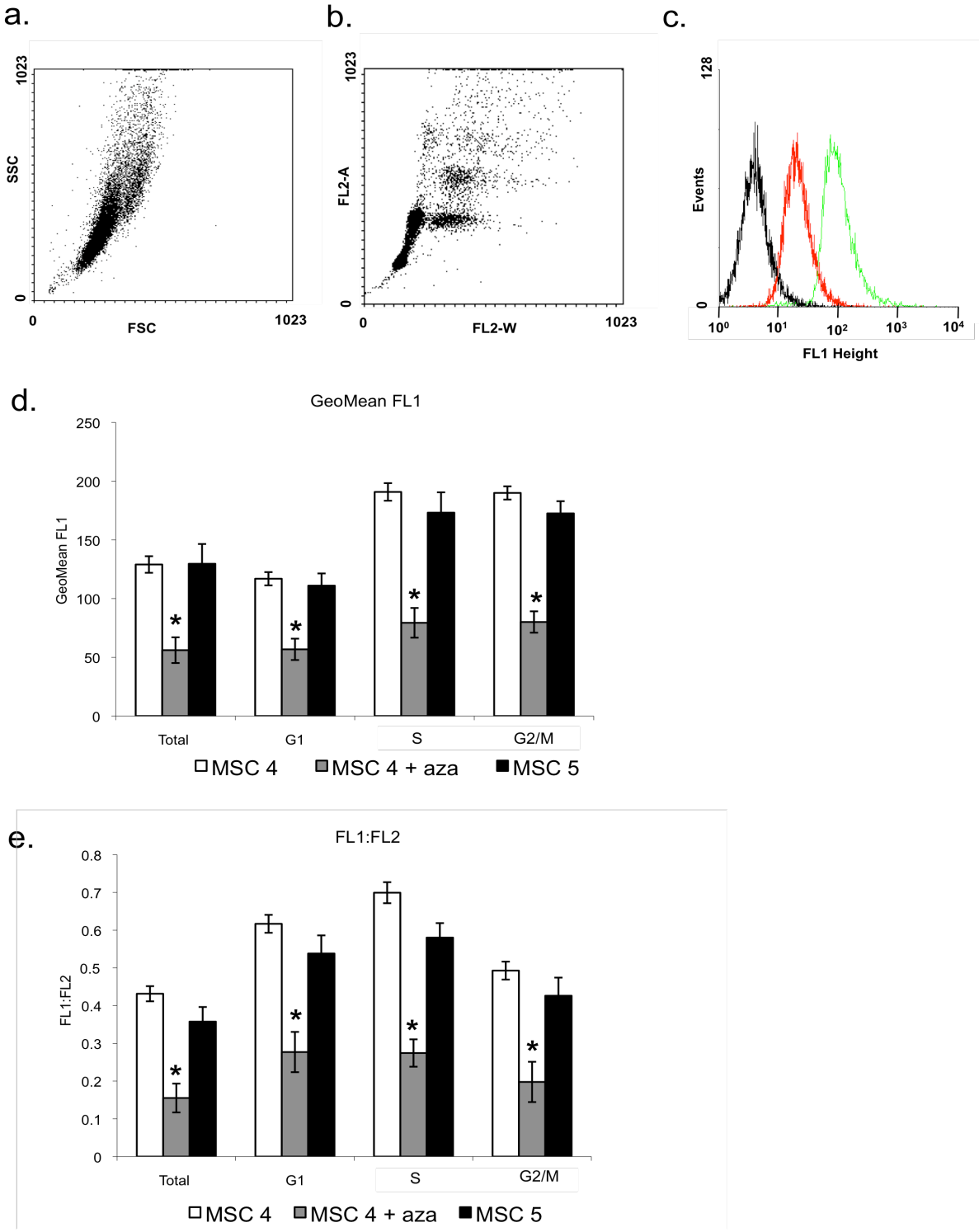


Figure 3.2: Analysis of global methylation by flow cytometry is unable to detect any significant difference in methylation between MSC 4 and MSC 5. Cells were stained with 0.05µg mouse anti-methylcytosine primary antibody and 8µg/ml anti-mouse FITC labelled secondary antibody and analysed by flow cytometry. **(a)** Dot plot displaying forward scatter (FSC) versus side scatter (SSC) for MSC 5 cells. **(b)** Dot plot displaying FL2 area (FL2-A) versus FL2 width (FL2-W) for MSC 5 cells, used for doublet discrimination and separation of

G1, M and G2 phase cells. **(c)** Histogram displaying FL1 signal of unstained (black), IgG1 isotype control stained (red) and anti-methylcytosine stained (green) MSC 5. **(d)** GeoMean FL1 measurements for total, G1 phase, S phase and G2/M phase populations of MSC 4, MSC 4 treated with 5μM 5-aza-deoxycytidine for four days (MSC 4 + aza) and MSC 5 cells. **(e)** FL1:FL2 ratio for total, G1 phase, S phase and G2/M phase populations of MSC 4, MSC 4 treated with 5μM 5-aza-deoxycytidine for four days (MSC 4 + aza) and MSC 5 cells. . Student's T-test was performed for all statistical analyses. * $p < 0.05$ wrt. MSC 4. Columns display average of three replicates; Error bars display SEM.

3.3.2 Global Methylation Analysis by MethyLight- Optimisation

In order to investigate repetitive element methylation in the MSC cell lines, the qRT-PCR based method MethyLight was utilized to assess methylation of Alu, pericentromeric satellite-2 (Sat2) and long interspersed nuclear element 1 (LINE1) repetitive elements. Methylation of these elements has previously been shown to correlate strongly with global methylation as measured by high performance liquid chromatography (Weisenberger et al., 2005). MethyLight reactions which showed the strongest correlation (Alu-M2, Sat2-M1 and LINE1-M1) were chosen for use in this study. The primer and probe sets for each reaction were designed to amplify methylated sequences from bisulphite converted DNA, with the efficiency of amplification depending on the degree of methylation present at the loci. An unbiased primer and probe set designed against a region of the consensus Alu sequence absent of CpG dinucleotides (Alu-C4) was utilized as a loading control and to control for efficiency of bisulphite conversion. All samples were run alongside fully methylated and unmethylated control DNA, with methylation levels of experimental samples expressed as a fold change relative to the fully methylated reference.

Optimisation of primer and probe concentrations were conducted for all MethyLight reactions to ensure maximal reaction efficiency (Figure 3.3). Starting concentrations for these optimisations were obtained from Weisenberger *et al.* and varied around these points. Primer optimisation was conducted by maintaining probe concentration at 0.1μM and varying primer concentration from 50-900nM, with several combinations of forward and reverse primer concentrations. The same template (fully methylated DNA) was used for each

reaction in equal amount (50ng). Interpretation of average threshold (C_t) values and standard error for these reactions indicated highest reaction efficiency at 900nM forward and reverse primers for all reactions (Figure 3.3 a-c).

Next, primer concentrations were maintained constant at 900nM and probe concentration was varied. As before, analysis of average C_t values and standard error indicated highest reaction efficiency at 0.2μM for Alu-M2 and LINE1-M1 reactions, and 0.1μM for the Sat2-M1 reaction (Figure 3.3 d-f).

To control for the efficiency of bisulphite conversion, all samples were treated in the same bisulphite reaction along with fully methylated (M.SssI treated) and unmethylated (GenomiPhi whole genome amplification product) controls. Every cell line was assayed in triplicate, with each PCR reaction also conducted in triplicate. The fully methylated and unmethylated reference samples were used as controls for the PCR reaction and as standards against which to express methylation levels. Final methylation values were expressed as a fold change relative to the fully methylated reference.

Figure 3.3

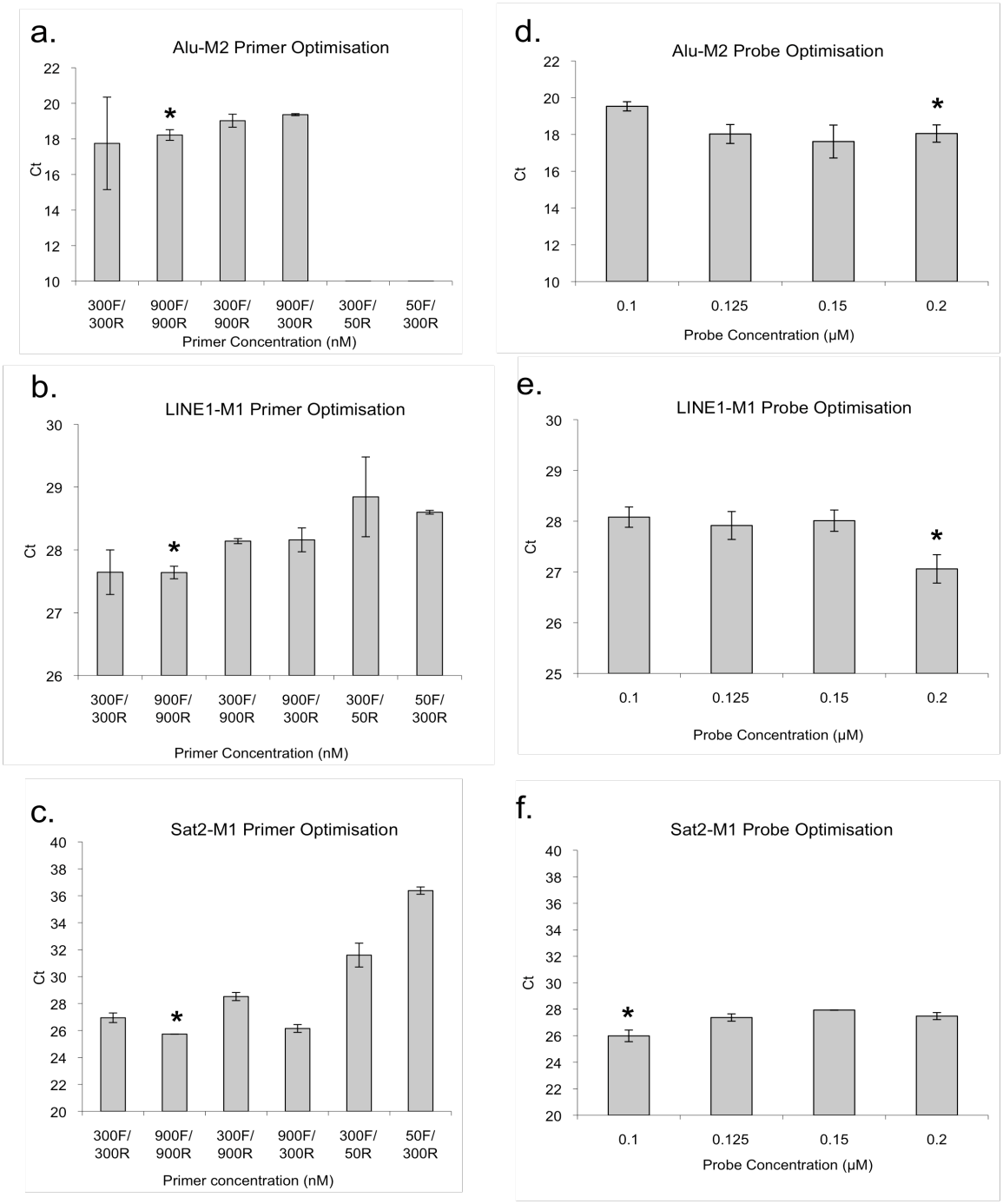


Figure 3.3: Optimisation of MethyLight Reactions. Quantitative PCR reactions using 50ng of fully methylated DNA as template were performed with varying concentrations of forward and reverse primers for the **(a)** Alu-M2, **(b)** LINE1-M1 and **(c)** Sat2-M1 MethyLight reactions, with average threshold (C_t) values recorded. Optimal primer concentrations were subsequently used for probe concentration optimisation for the same three reactions **(d-f)**. MethyLight reactions were performed as before but with probe concentration was varied. Columns display average of three replicates; Error bars display SEM. * optimal concentration used in further experiments.

3.3.3 5-aza-deoxycytidine Treatment of MSC and HF

To test the sensitivity of the MethyLight reactions, MSC and HF cells were treated with 5-aza-deoxycytidine for four days in order to induce global hypomethylation (Figure 3.4). MethyLight analysis of Alu, Sat2 and LINE1 elements in MSC 0 demonstrated a decrease in methylation with increasing concentrations of 5-aza-deoxycytidine. Transformed MSC 5 cells, however, did not respond well to the treatment, displaying increased cell death alongside unchanged or increased methylation values as determined by MethyLight (Figure 3.4 a-c). Hypomethylation of all elements was observed in parental and transformed HF treated with 1μM 5-aza-deoxycytidine for four days (Figure 3.4 d-f). These results indicated that MethyLight was able to detect global hypomethylation at these repetitive elements, and that hypomethylation could be successfully induced in parental mesenchymal stem cells and human fibroblasts.

Figure 3.4

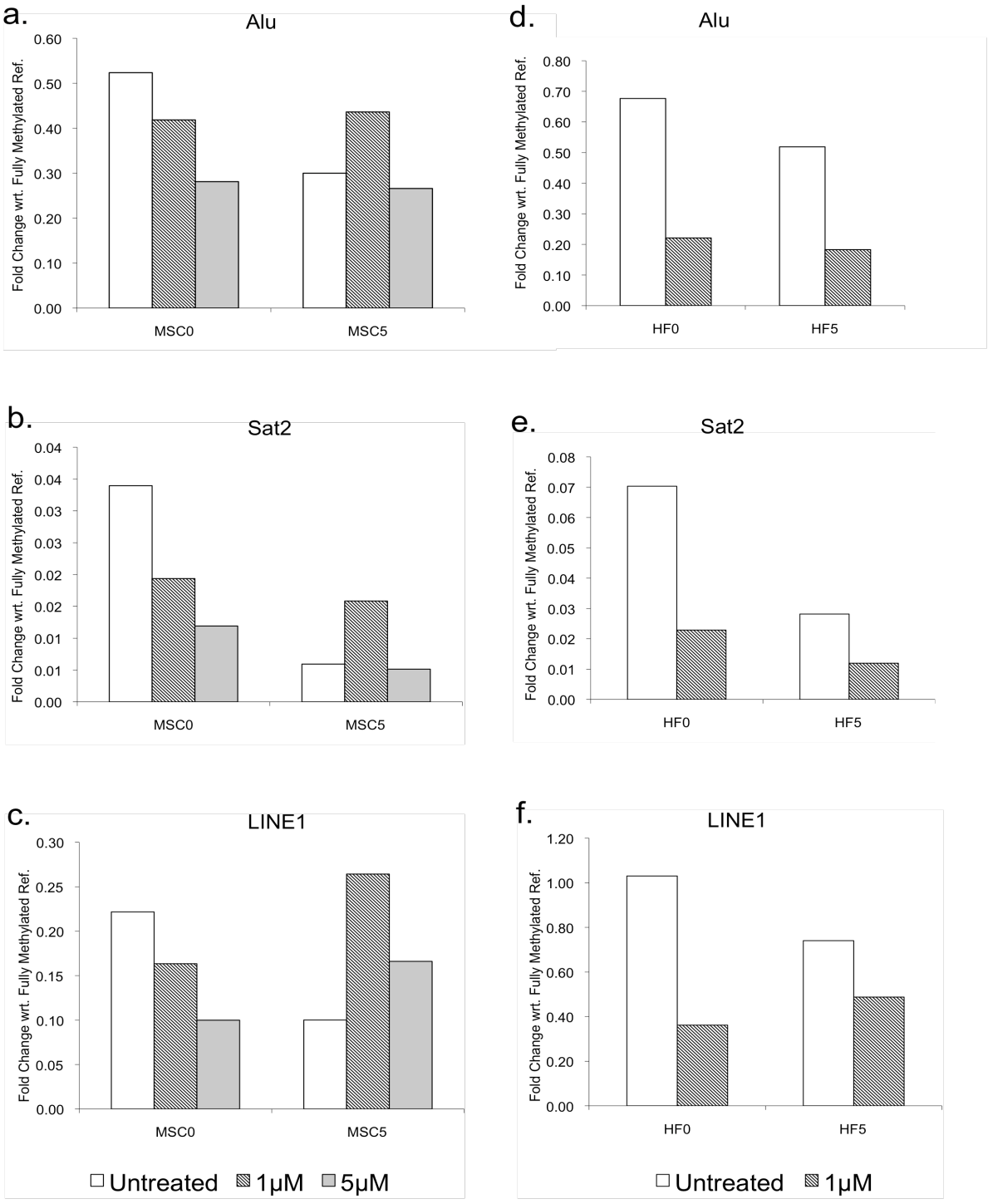


Figure 3.4: 5-aza-deoxycytidine treatment induces repetitive element hypomethylation in MSC and HF that is detectable by MethyLight. Parental MSC (MSC 0), transformed MSC (MSC 5), parental HF (HF 0) and transformed HF (HF 5) were treated with increasing concentrations of 5-aza-deoxycytidine for 4 days. Fold change in methylation of **(a)** Alu, **(b)** Sat2 and **(c)** LINE1 elements in MSC and **(d)** Alu, **(e)** Sat2 and **(f)** LINE1 in HF with respect to fully methylated reference DNA was assayed by MethyLight as described previously.

3.3.4 Analysis of Global Methylation levels in the MSC Model by MethyLight

To determine whether genome-wide hypomethylation takes place during step-wise transformation of MSC, methylation of Alu, Sat2 and LINE1 repetitive elements was compared between MSC 0 and MSC 5 by MethyLight. These data showed no change in Alu methylation but a significant decrease in Sat2 and LINE1 methylation between these two cell lines (Figure 3.5 a-c).

To identify the step at which this decrease in methylation occurs during transformation, Sat2 methylation was assessed in more detail at each step during transformation of MSC (Figure 3.5d). This element was chosen for further investigation because it provided a strong indicator of global methylation levels and also displayed the most significant decrease in methylation in this model. These data indicated a decrease in Sat2 methylation between MSC 0 and MSC 1 (on expression of hTERT), followed by a return to similar levels with p53 inhibition in MSC 2E6. With further genetic hits, Sat2 methylation remained similar to parental MSC levels until the introduction of H-Ras^{V12} between MSC 4 and MSC 5, at which point a significant decrease was observed. Sat2 hypomethylation was also maintained in MSC 5 tumours grown in mice (Figure 3.5d).

To test whether these observed changes were specific to the mesenchymal stem cell model, MethyLight analysis was also conducted in human fibroblasts (HF) transformed using the same 5 oncogenic steps and in tumour derived by injection of transformed HF into nude mice. No decrease in methylation was observed for any of the repetitive elements assayed (Figure 3.6)

Figure 3.5

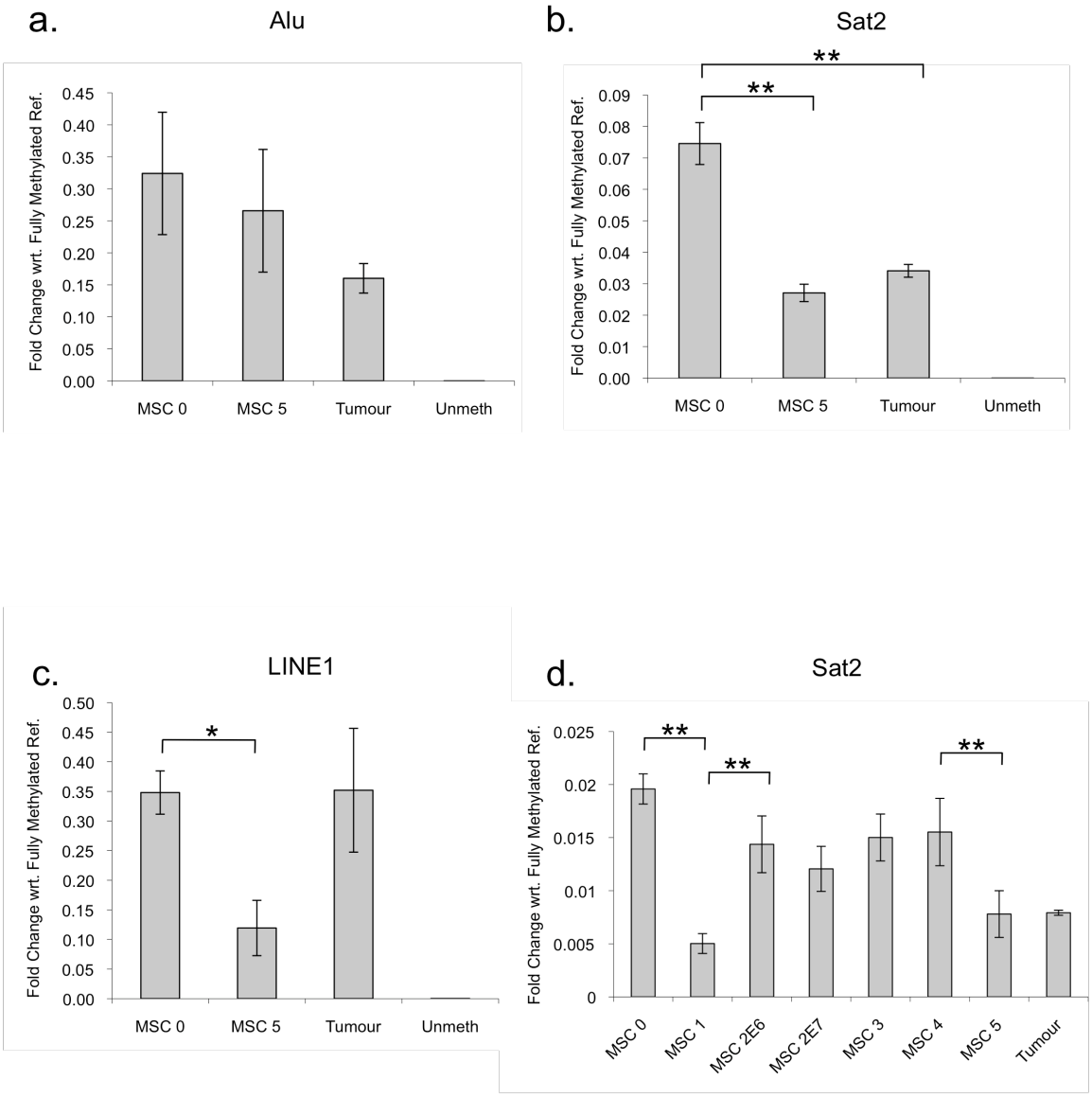


Figure 3.5: MethyLight analysis of repetitive element methylation shows decrease in Sat2 and LINE1 methylation between parental and transformed MSC. Quantitative PCR was conducted using primers against methylated **(a)** Alu, **(b & d)** Sat2 and **(c)** LINE1 sequences in parental MSC (MSC 0), transformed MSC (MSC 5), tumours derived by injecting transformed MSC into nude mice (Tumour) and fully unmethylated GenomiPhi amplified DNA (Unmeth). **(d)** Quantitative PCR for the Sat2 element was also conducted in all MSC cell lines (MSC 0 – 5) and tumours. Fold-change in methylation was calculated in three replicates by ΔΔCt using an unbiased Alu reaction as a loading control and normalised to fully methylated (M.SssI treated) reference DNA. Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$. Columns display average of three replicates; Error bars display SEM.

Figure 3.6

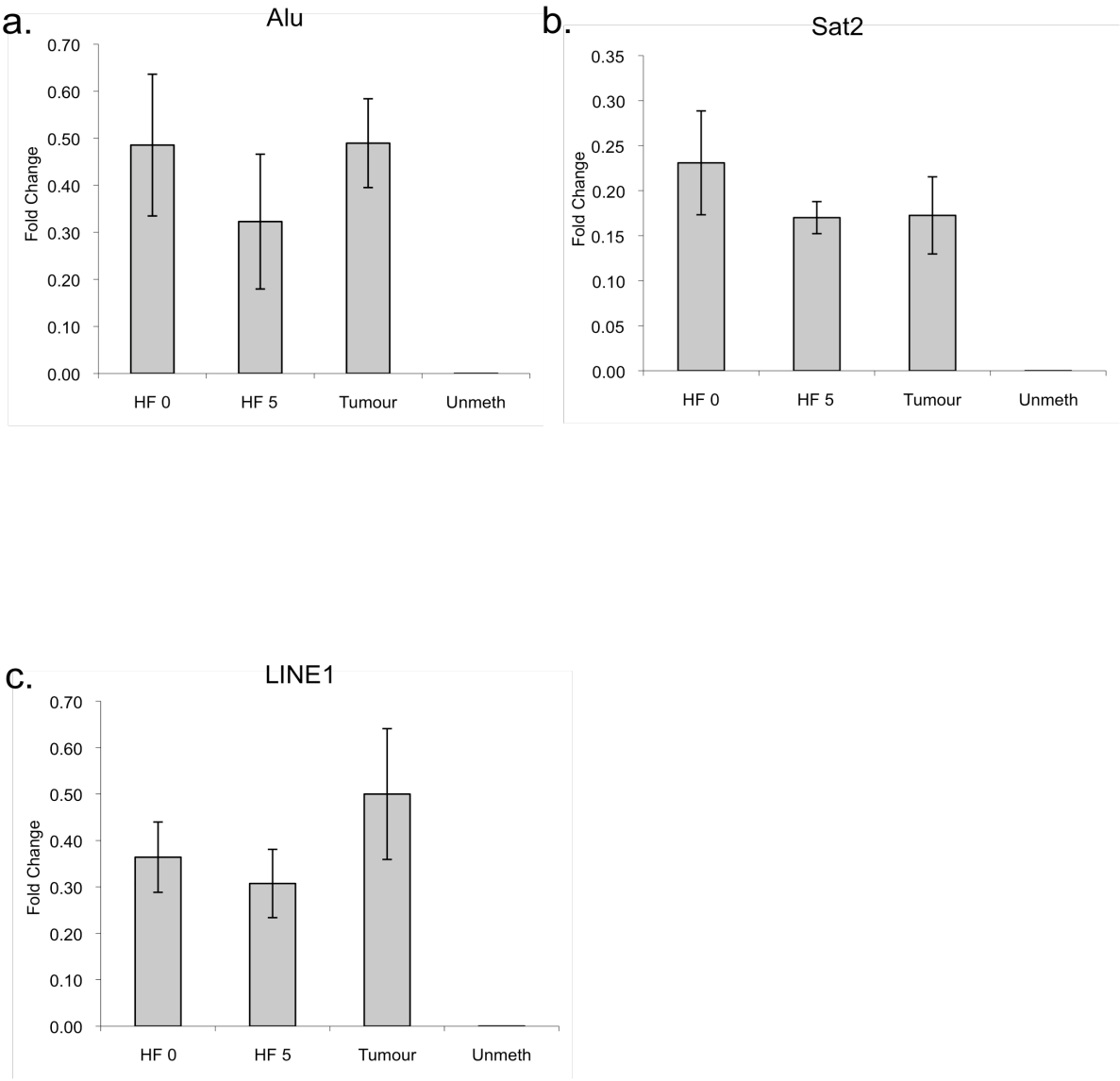


Figure 3.6: MethyLight analysis of repetitive element methylation shows no change in methylation between parental and transformed HF. Quantitative PCR was conducted using primers against methylated **(a)** Alu, **(b)** Sat2 and **(c)** LINE1 sequences in parental HF (HF 0), transformed HF (HF 5), tumours derived by injecting transformed HF into nude mice (Tumour) and fully unmethylated GenomiPhi amplified DNA (Unmeth). Fold-change in methylation was calculated in three replicates by ΔΔCt using an unbiased Alu reaction as a loading control and normalised to fully methylated (M.SssI treated) reference DNA. Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$. Columns display average of three replicates; Error bars display SEM.

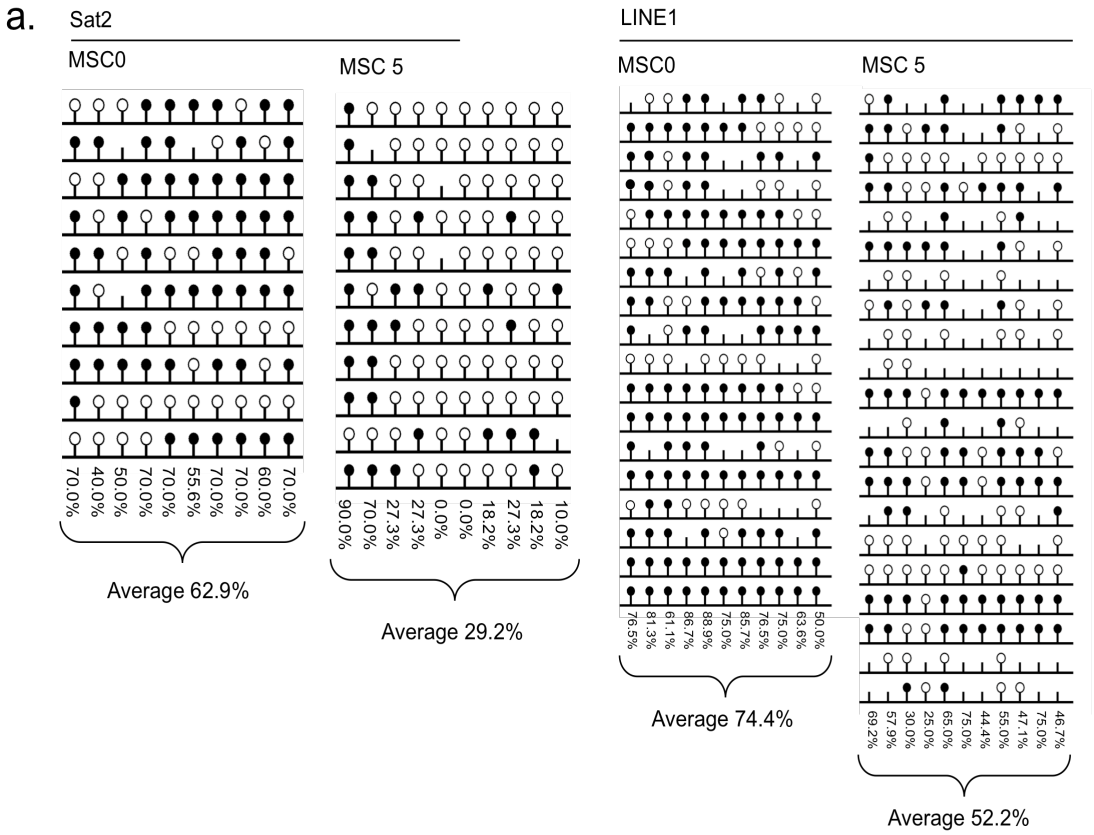
3.3.5 Bisulphite Sequencing Confirmation of MethyLight Data

To confirm the results observed by MethyLight, bisulphite sequencing of the Alu, Sat2 and LINE1 elements was conducted (Figure 3.7). Control primers without a bias for methylated or unmethylated sequences were designed against the regions assayed previously by MethyLight. These regions were amplified by PCR and cloned into the pGEM-T vector, with 10-25 individual clones sequenced in order to assess CpG methylation across the element. The PCR products covered 3 CpG dinucleotides in the Alu element, 11 in LINE1 and 10 in Sat2. Analysis of parental and transformed MSC and HF confirmed the MethyLight data, demonstrating a significant decrease in both Sat2 and LINE1 element methylation between MSC 0 and 5, but no change in Alu methylation (Figure 3.7 a-d). Further analysis was conducted to confirm that this decrease in Sat2 and LINE1 methylation occurred on the introduction of H-Ras^{V12}. Bisulphite sequencing of Sat2 and LINE1 elements was conducted in MSC 4 and 5, indicating that hypomethylation of these elements was induced on H-Ras^{V12} overexpression. No changes in Alu, Sat2 or LINE1 methylation were observed between HF 0 and 5, confirming the data obtained by MethyLight (Figure 3.7 a-d).

To establish whether hypomethylation of these elements occurred in a site specific manner, average methylation values for each CpG site were compared between parental and transformed MSC and HF (Figure 3.8). This indicated a drastic decline in methylation across CpG sites 5-7 in the Sat2 element between parental and transformed MSC and HF. Analysis of this region for transcription factor binding sites using TF Search

(http://www.cbrc.jp/research/db/TFSEARCH.html) revealed potential binding sites for v-Myb and c-Ets in this region (Figure 3.8b).

Figure 3.7



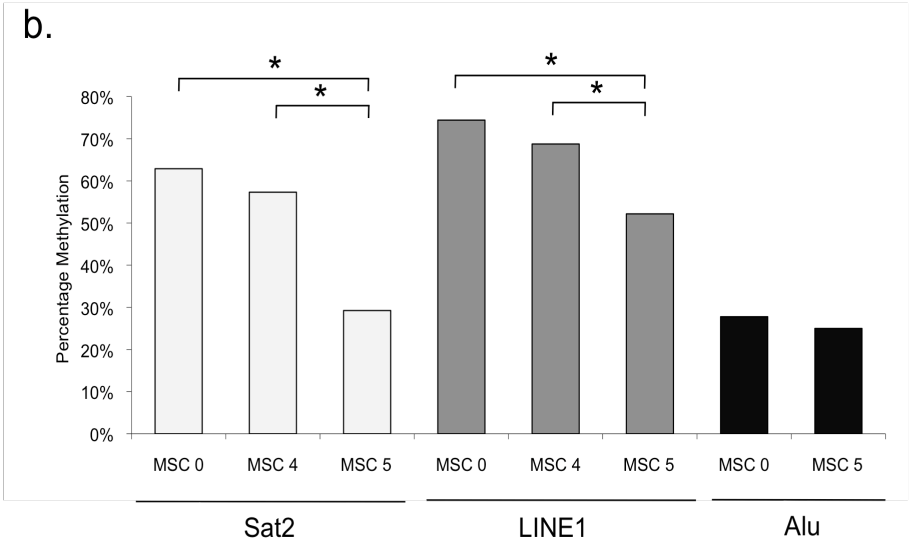


Figure continued on the following page

Figure 3.7

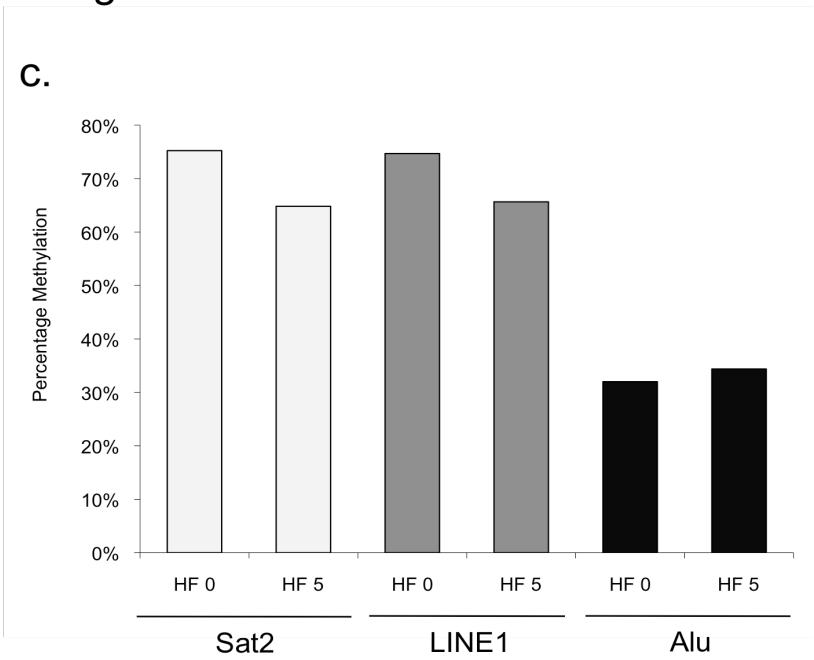
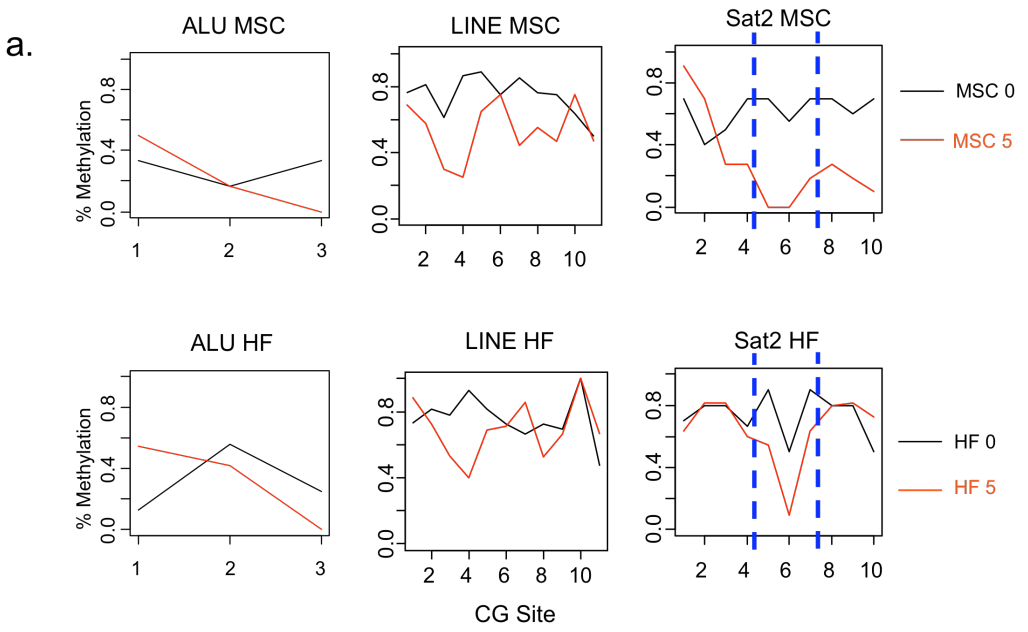


Figure 3.7: Bisulphite sequencing of repetitive elements confirms MethyLight data. Sat2 and LINE1 elements were amplified by PCR from bisulphite modified genomic DNA with individual PCR products cloned into pGEM-T and sequenced. **(a)** Representative sequencing data for Sat2 and LINE1 elements, with each line displaying one cloned PCR product. Open circles represent unmethylated CpG sites, closed circles methylated CpGs and absent circles polymorphic sites where the expected CpG is not present. Average methylation is also expressed as a percentage of all CpG sites assayed for **(b)** Sat2, **(c)** LINE1 **(d)** and Alu elements.

Figure 3.8



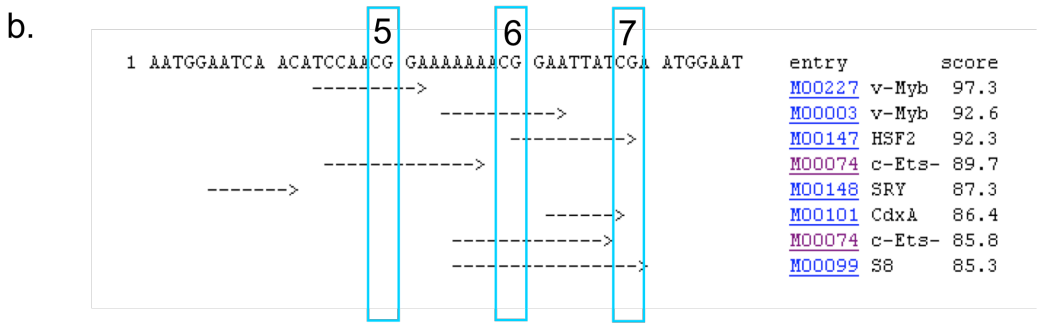


Figure 3.8: Sat2 hypomethylation occurs in a site specific manner. (a) Average methylation values displayed for each CpG site across the Alu, Sat2 and LINE1 repetitive elements. Blue lines indicate the region of Sat2 at which the greatest loss of methylation is observed between MSC 0 and MSC 5. **(b)** Transcription factor binding site analysis of Sat2 element CpG sites 5-7. * p <0.05 Paired Wilcoxon Signed Rank Test. Columns display average methylation percentages.

3.3.6 *Isolation and Characterisation of Clonal MSC5 Populations*

The MSC 5 cell line was initially generated as a mixed population following hygromycin selection of MSC 4 cells infected with H-Ras^{V12} retrovirus. In order to further characterize this mixed population, and investigate whether global hypomethylation occurs within a specific sub-population of cells, clonal MSC 5 populations were isolated (Figure 3.9). MSC 5 were plated at a density of 100 cells per 10cm dish and grown until individual colonies were visible. These colonies were individually trypsinised, replated and expanded. Following isolation, individual clones were shown to express different levels of H-Ras^{V12}, with concurrent variation in downsteam MAPK signalling as measured by levels of phospho-ERK1/2 (pERK) relative to total ERK1/2 (Figure 3.9a). In order to confirm that the clonal cell lines were transformed, *in vitro* growth of the cells in soft agarose was assessed. All cell lines were able to form colonies in soft agarose, although a significant variation in the average colony size after 12 days growth was observed (Figure 3.9b). Comparison of average colony diameter to H-Ras protein levels indicated a strong correlation between the two ($r^2 = 0.72$; Figure 3.9c). To determine whether increased H-Ras^{V12} expression correlated with greater hypomethylation in the MSC 5 clones, bisulphite sequencing of the Sat2 element was conducted in low H-Ras (MSC 5, clone 2) and high H-Ras (MSC 5, clone 5) expressing populations. This indicated that Sat2 elements in both high and low H-Ras expressing lines were hypomethylated compared to MSC 4. Sat2 methylation in clone 5 (high H-Ras) was lower than in clone 2 (low H-Ras), although this difference was not statistically significant (Figure 3.9d). These data show that both high and low H-Ras expressing clones within the mixed MSC 5 population display hypomethylation of the Sat2 pericentromeric

repeat, although the degree of hypomethylation did not significantly correlate with H-Ras expression.

Figure 3.9

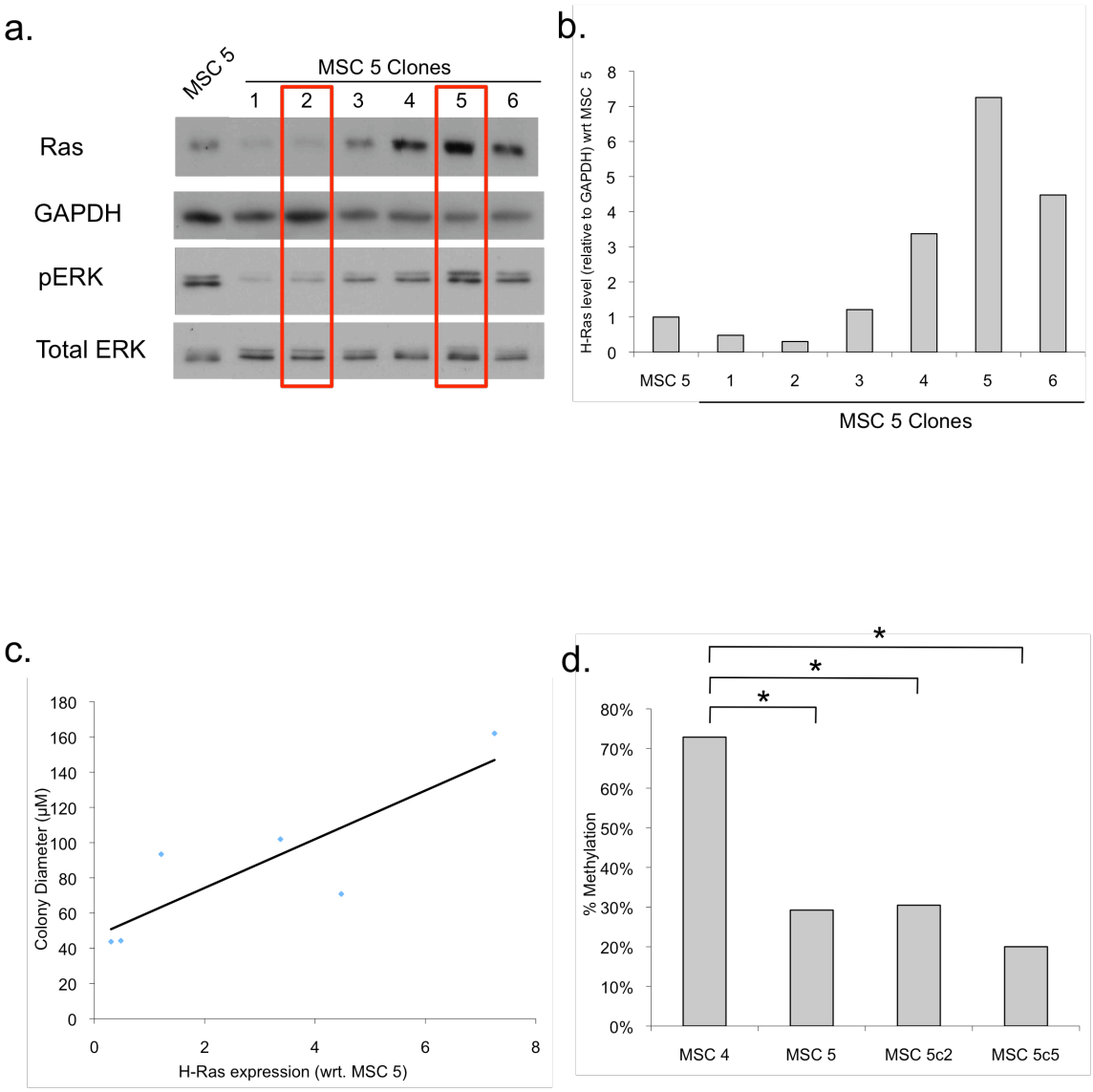


Figure 3.9: H-Ras expression and repetitive element methylation in MSC 5 clones. **(a)** Western blot showing Ras, GAPDH, phosphorylated MAPK1/2 (pERK) and total MAPK1/2 (total ERK) levels in the original MSC 5 population (MSC 5) and six clonal MSC 5 populations. **(b)** Quantification of H-Ras levels relative to the original MSC 5 population by densitometry. **(c)** Correlation between H-Ras levels and colony diameter in soft agarose transformation assay. **(d)** Percentage methylation of Sat2 repetitive element as determined by bisulphite sequencing in MSC 4, MSC 5, a low H-Ras expressing clone (MSC 5 Clone 2) and a high H-Ras expressing clone (MSC 5 Clone 5). * $p < 0.05$ Paired Wilcoxon Signed Rank Test. Columns display averages.

3.3.7 MSC 5 Karyotype

Loss of methylation at repetitive elements can lead to increased genomic instability and abnormal karyotypic changes. To establish whether such changes occur during step-wise transformation of MSC, a full karyotypic analysis of the mixed MSC 5 population including multicolour fluorescence *in situ* hybridisation (MFISH) was conducted (Figure 3.10). This revealed that the MSC 5 population consisted of several sub-populations with unique translocations, including a tetraploid sub-population making up approximately 17% of the total number of cells (Figure 3.10c). The fully karyotype of this cell line is as follows:

46,XY,der(9)t(9;?12)(q34;?),der(10)t(10;12)(p1?1;?),der(21) t(?12;21)(?;p10) [cp16]/47,idem,+der(?20)t(20;21)(p1?1.2;?) [2]/92,XXYY,idemx2,inc[2]

Further analysis of cell ploidy by FACS revealed that this tetraploid population was also present in the MSC 4 population at a level of approximately 10% (Figure 3.10d).

Figure 3.10

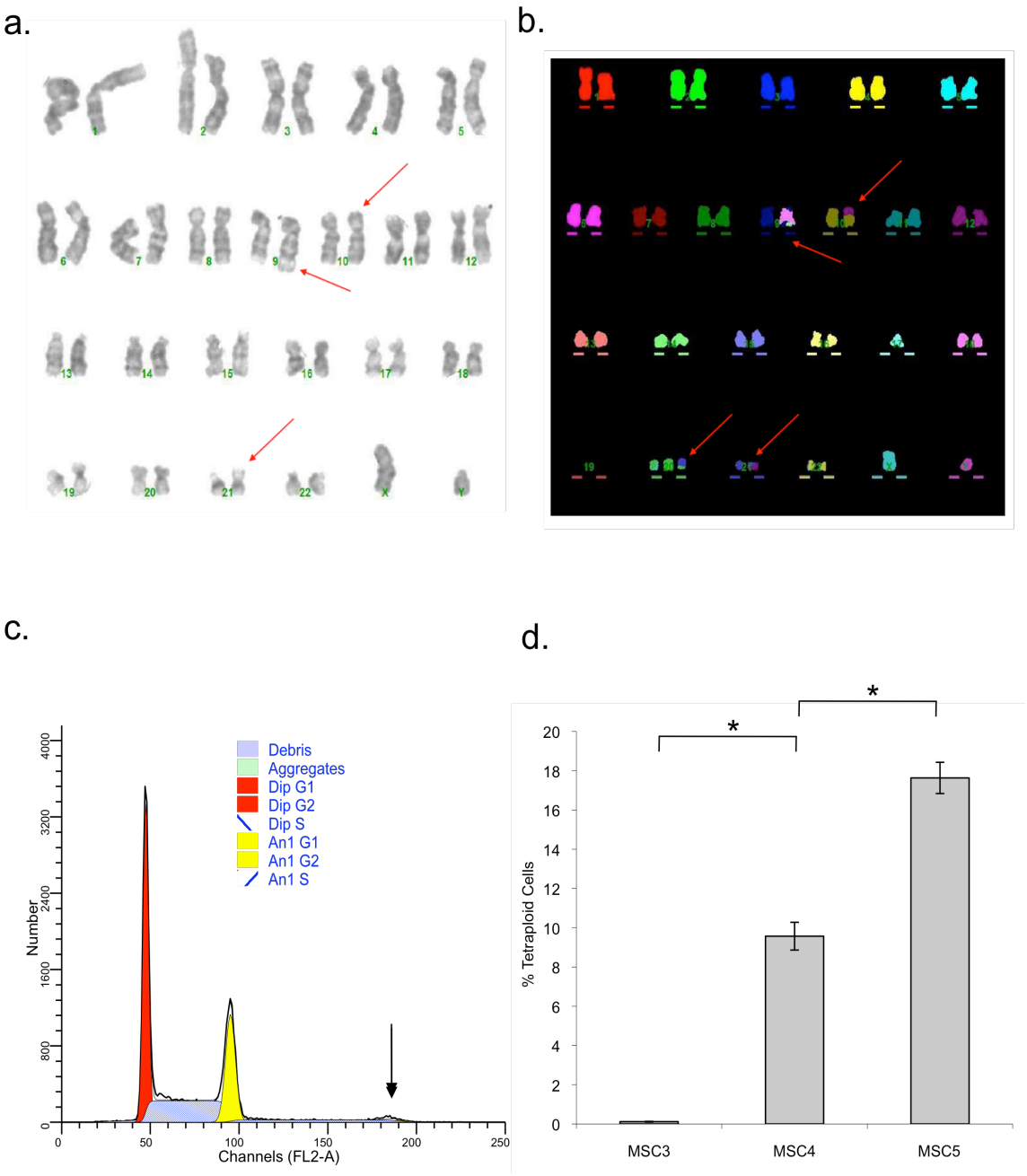


Figure 3.10: MSC 5 karyotype. **(a)** Karyogram showing, arrowed, der(9), der(10) and der(21). **(b)** Partial side line MFISH karyogram showing, arrowed, der(9), der(10), der(20) and der(21). **(c)** Cell cycle analysis of MSC 5. Cells were grown as normal, fixed and permeabilised, stained with propidium iodide and analysed by flow cytometry. Tetraploid G2 population is arrowed. **(d)** Percentage of tetraploid cells in the MSC 3, 4 and 5 populations were calculated using the ModFit software package. Student's T-test was used for all statistical analyses. * $p < 0.05$. Columns display average of three replicates; Error bars display standard deviation.

3.4 Discussion

Global hypomethylation has been observed in a wide range of cancers as assayed by a number of different techniques, many of which utilize repetitive elements as markers for genome wide methylation status. Despite the wealth of studies on global DNA hypomethylation, very little progress has been made to elucidate the mechanism of this loss of methylation.

Our laboratory developed a stem cell model of cancer in which mesenchymal stem cells (MSC) have been transformed through 5 oncogenic steps (hits) – the introduction of hTERT, HPV E6 and E7, SV40 small T and oncogenic H-Ras (Funes et al., 2007). This stepwise model provides a well defined genetic background in which to investigate epigenetic changes that occur during the process of transformation. Studies of the methylation status in this model could therefore provide indications of the causes of global hypomethylation in cancer.

My studies on this model have revealed a number of interesting findings. Firstly, hypomethylation of the Sat2 and LINE1 repetitive elements is observed during transformation, with the decrease in methylation occurring on the introduction of oncogenic H-Ras^{V12} in the final step. Hypomethylation of these elements has been widely observed and shown to correlate strongly with genome-wide methylation levels, suggesting that genome-wide hypomethylation occurs in this system as a result of H-Ras^{V12} expression in MSC 5 (Ehrlich, 2002; Weisenberger et al., 2005). Although these repetitive elements constitute a sizable percentage of the genome, hypomethylation was not observed in the most common repeat, the Alu element, which alone makes up approximately 10% of

the human genome (Lander et al., 2001). Few studies have assessed changes in Alu methylation in cancer, with those that have producing contradictory results. One study of neuroendocrine tumours noted reduced Alu methylation in this cancer compared to adjacent normal tissue (Choi et al., 2007). Similarly, Alu hypomethylation has been detected in the chronic phase of chronic myeloid leukaemia (CML) and blast crisis CML compared to normal bone marrow cells (Roman-Gomez et al., 2008). However, a study of seminoma and dysgerminoma observed an increase in Alu methylation when compared to normal sperm and spleen, although this may be complicated by the hypomethylated state of Alu repeats in normal sperm (Rubin et al., 1994). Finally, a genome-wide study of repetitive element methylation by MeDIP and high throughput sequencing in malignant peripheral nerve sheath tumours (MPNST) noted similar patterns of hypomethylation to those observed in my investigation (Figure 3.11). Results from this report indicate that satellite repeats, and to a lesser extent LINE elements, are hypomethylated in the cancer compared to normal control tissue, whereas SINE elements (including Alu repeats) show no change in methylation (Feber *et al.*, in press). It is possible, therefore, that certain repetitive elements, such as satellite and LINE sequences, may be more likely to display DNA hypomethylation in cancer.

Although I was able to detect loss of methylation in the Sat2 and LINE1 elements by bisulphite sequencing and MethyLight, this difference was not apparent when global methylcytosine levels were assayed by FACS. Antibody optimization curves and matched IgG1 controls indicated that staining is specific for methylcytosine. Furthermore, a significant decrease in the methylcytosine signal is observed in MSC 4 cells treated with 5-aza-deoxycytidine to induce

hypomethylation, suggesting that changes in genome-wide methylation levels are detectable by this method. Although there is a small decline in the amount of methylcytosine relative to total DNA (FL1:FL2 ratio) in MSC 5 compared to MSC 4, this change does not achieve statistical significance. Therefore, the hypomethylation of Sat2 and LINE1 elements in MSC 5 may not constitute a large enough demethylation event to be detectable using this method.

The widespread incidence of genome wide hypomethylation in a range of cancers, including benign neoplasms, has suggested that this epigenetic change occurs early during the process of transformation (Feinberg et al., 2006; Goelz et al., 1985; Suter et al., 2004). The results obtained from this step-wise model of transformation are contrary to these studies, indicating that hypomethylation occurs in fully transformed cells following the final oncogenic 'hit'. This suggests that hypomethylation is a late event during step-wise transformation, supporting previous reports that hypomethylation arises in the latter stages of cancer progression (Gama-Sosa et al., 1983a; Yegnasubramanian et al., 2008). However, transformation in the MSC model is induced by a series of genetic hits and may not fully recapitulate the true early epigenetic changes that occur during *in vivo* transformation as suggested by the epigenetic progenitor model of cancer (Feinberg et al., 2006). It is possible, therefore, that although global hypomethylation is a late event in this model, it may occur earlier during *in vivo* cancer progression as a prelude for further genetic and epigenetic changes.

Previous reports have suggested links between Ras signalling and changes in DNA methylation, including incidences of Ras induced hypermethylation and hypomethylation. These include hypermethylation of a

CpG island upsteam of the clusterin promoter and a CpG island of the MMP2 promoter in H-Ras transformed rat fibroblasts (Lund et al., 2006). Additionally, genome-wide methylation levels in the adrenocortical tumour cell line Y1 have been shown to be regulated by Ras mediated modulation of DNA methyltransferase activity, with reduced Ras signalling resulting in genome-wide hypomethylation (MacLeod et al., 1995; Rouleau et al., 1995). One of the most well characterised examples of Ras dependent hypermethylation is the down-regulation of the pro-apoptotic *Fas* gene in NIH 3T3 cells. Investigation of this system revealed a range of factors required for Ras-mediated epigenetic silencing of *Fas* including DNMT1, EZH2, EED and HDAC9 (Gazin et al., 2007). As well as this role in DNA hypermethylation, studies have suggested that Ras may also be involved in the induction of DNA hypomethylation. For example, my findings are consistent with a study of v-H-Ras induced general DNA demethylation in mouse embryonal P19 cells (Szyf et al., 1995). Although these previous studies demonstrate that there is an interaction between Ras signaling and the control of DNA methylation in cells, my investigation is the first to link Ras signaling and global hypomethylation during step-wise transformation in human cells.

In this study, I observed hypomethylation in MSC transformed using 5 oncogenic hits. As a control for cell type, I conducted parallel analyses in a model of step-wise transformation whereby differentiated human fibroblasts were transformed using the same oncogenic hits. These cells, however, did not exhibit hypomethylation of any repetitive elements tested. It is possible that differences in the epigenetic state of pluripotent MSC and differentiated HF account for these differences. Chromatin modifications and DNA methylation have been profiled on a genome wide scale in pluripotent and differentiated cells

(Azuara et al., 2006; Meissner et al., 2008; Mikkelsen et al., 2007). Interestingly, pluripotent cells display a characteristic epigenetic signature, with the promoters of key developmental regulators in a 'bilvalent' state – repressed with the potential to be activated during lineage commitment. Furthermore, DNA methylation patterns strongly correlate with histone methylation patterns and undergo changes during differentiation. Stem cells therefore maintain pluripotency and initiate different patterns of differentiation through a greater epigenetic 'plasticity' than differentiated cells committed to a particular epigenetic state. It is likely, therefore, that the initial epigenetic background of the MSC and HF and the ability to change this epigenetic patterning were radically different at the start of the transformation process. Human fibroblasts, therefore, may lack the epigenetic plasticity required for this loss of methylation to take place.

Both low and high H-Ras expressing clones were isolated and shown to exhibit hypomethylation of Sat2 elements compared to MSC 4, although there was no statistically significant correlation between the degree of Sat2 hypomethylation and H-Ras expression. This suggests that Sat2 hypomethylation in these cells was dependent on transformation rather than H-Ras expression levels. Additionally, hypomethylation of Sat2 elements can be observed in at least two of the clonal cell populations, suggesting that a single sub-population of cells is not responsible for the hypomethylation observed in the original MSC 5 population.

Our studies of the MSC 5 karyotype also confirmed that a number of distinct cell populations displaying abnormal recombination events were present in the mixed MSC population. This is consistent with the hypothesis that

hypomethylation of repetitive elements can lead to increased genomic instability and abnormal karyotypic changes. For example, ICF syndrome is characterized by hypomethylation of Sat2 elements, resulting in specific recombination events at genomic regions where these elements are located on chromosomes 1 and 16 (Tuck-Muller et al., 2000). Similarly, a correlation between LINE1 and Sat-α hypomethylation and karyotypic changes has been observed in multiple myeloma (Bollati et al., 2009). Although the MSC 5 cell line does display a number of recombination events, none appear to involve recombination events between chromosomes 1 and 16 as observed in ICF. Our karyotype analysis also indicated the presence of a tetraploid sub-population making up approximately 17% of the total MSC 5 population. Further characterization of the percentage of tetraploid cells by FACS indicated that this population was also present in MSC 4, albeit to a smaller degree (~10% of cells). Although the link between abnormal cell ploidy and genome-wide hypomethylation has not been extensively investigated, one study has shown that LINE1 and Sat-α elements are less methylated in non-hyperdiploid compared to hyperdiploid multiple myelomas (Bollati et al., 2009). Additional karyotypic studies of the other MSC cell lines are required to determine whether these abnormalities are a direct result of repetitive element hypomethylation in MSC 5.

An important consideration for the interpretation of these findings is the cell cycle length at each stage of this model. The typical seeding dilutions for each MSC cell line indicate a significant shortening of the cell cycle with increasing genetic 'hits'. This has important implications for studies of methylation in these cell populations, since the proportion of cells in S-phase (containing hemi-methylated DNA) will differ between cell lines. The effect of

cell cycle length on repetitive element DNA methylation status can be observed in my studies of 5-azacytidine treatment of MSC 5. The increase in Alu, Sat2 and LINE1 methylation levels on 1uM 5-azacytidine treatment could result from slower cell growth, decreasing the proportion of cells in S-phase compared to the untreated population. This would in turn reduce the proportion of cells containing hemi-methylated DNA, leading to the observed increase in methylation levels that are overcome by the greater demethylating effect of drug treatment at higher concentration. Differences in cell cycle length between these cell lines must therefore be taken into consideration when interpreting observations in this model.

In conclusion, this study has demonstrated a link between H-Ras^{V12} expression and hypomethylation of repetitive elements during a step-wise model of transformation in human MSC. This model provides a controlled background in which known genetic hits can be linked to observed epigenetic changes. By transforming cells in this manner, I am able to compare transformed cells to the most appropriate normal cell controls, therefore overcoming problems arising from variation in methylation between cell types. Our results provide the first step towards elucidating a mechanistic basis for genome wide hypomethylation in cancer and a valuable model for further establishing the consequences of this change.

Figure 3.11

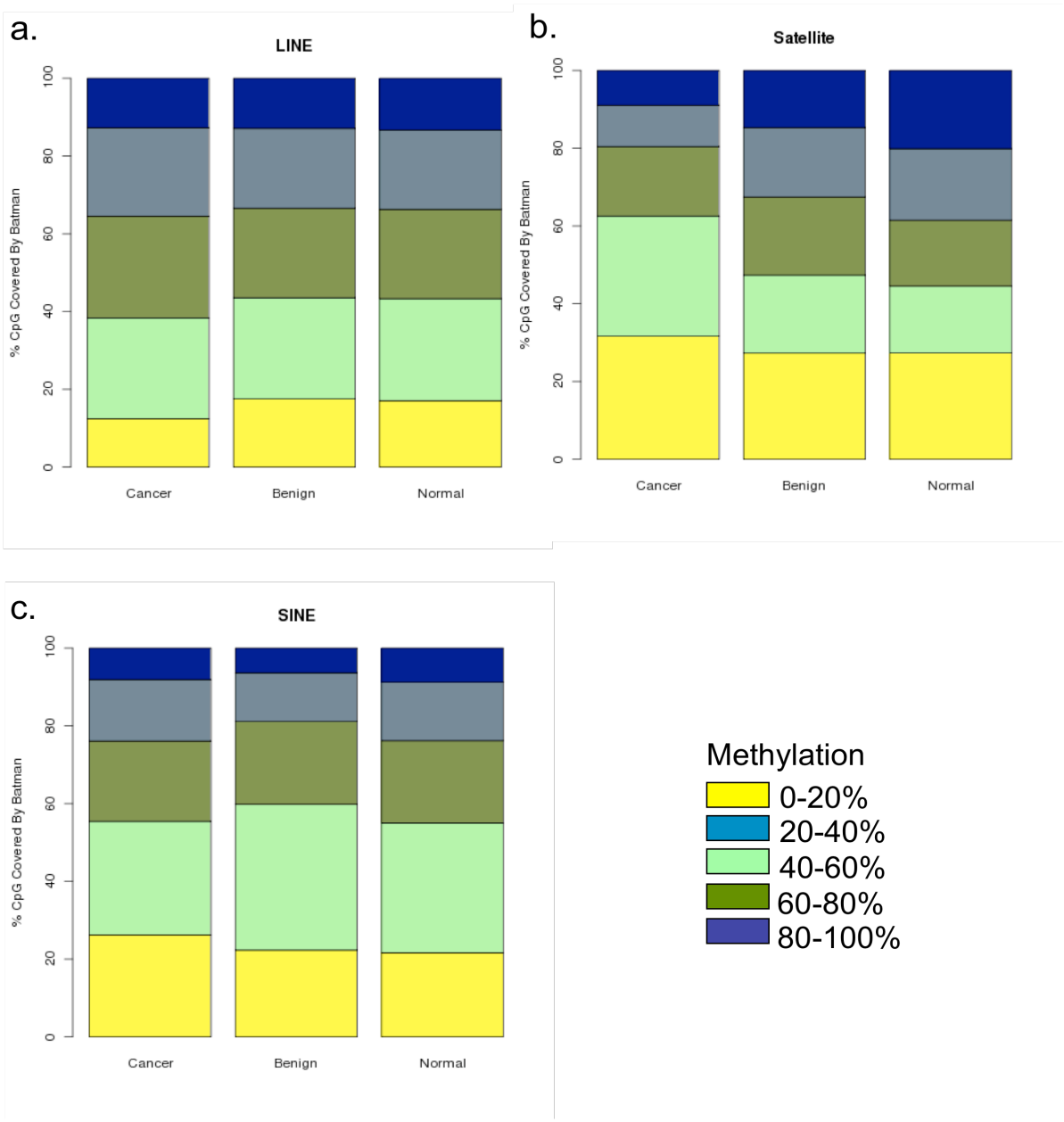


Figure 3.11: Analysis of repetitive element methylation in MPNST by ChIP-seq. Global analysis of repeat family methylation suggests that there is no overall change in SINE methylation in MPNST (cancer) compared to NF (benign) or normal Schwann cell tissue. However, microsatellites display significant hypomethylation in the cancer compared to benign and normal tissues. A small increase in LINE hypomethylation is also evident. Adapted from Feber *et al*. (in press).

Chapter 4: Genome-wide hypomethylation occurs gradually and is not essential for transformation of mesenchymal stem cells

4.1 Background

Although the incidence and consequences of genome-wide hypomethylation in cancer have been extensively studied, relatively little remains known about the causes of this loss of methylation. DNA methylation patterns result from competing processes that establish, maintain and remove this epigenetic mark. Any shift in the balance of these processes could therefore result in genome wide alterations to DNA methylation such as genome-wide hypomethylation. Specifically, genome-wide hypomethylation may occur in a 'passive' manner through inefficient DNA methylation maintenance, or in an 'active' manner by direct enzymatic removal of methyl groups from DNA.

The importance of DNA methyltransferases (DNMTs) in the establishment and maintenance of DNA methylation suggests that disruption of normal DNMT activity or expression could lead to the genome-wide DNA hypomethylation observed in cancer. For example, loss of the maintenance methylation activity of DNMT1 could result in passive DNA demethylation through consecutive rounds of DNA replication. In line with this, *dnmt1* knockout mice display a loss of ~90% of genome-wide methylation levels, resulting in embryonic lethality (Li et al., 1992). Conditional *dnmt1* knockout in mouse fibroblasts also causes genome-wide demethylation and p53 mediated cell death (Jackson-Grusby et al., 2001). Similar effects have been observed in

human colon cancer cell lines in which *DNMT1* can be knocked out in an inducible manner, leading to cell cycle arrest, global genetic instability and apoptosis as a result of genome-wide loss of DNA methylation (Chen et al., 2007). Additionally, disruption of *de-novo* methyltransferase expression could also induce global hypomethylation, as observed in *dnmt3b* knockout mouse embryonic fibroblasts (Dodge et al., 2005) and *DNMT3B* mutations in ICF patients (Hansen et al., 1999; Xu et al., 1999).

As well as passive demethylation processes, global hypomethylation could be induced in an active manner through increased expression or activity of DNA demethylases (Ooi and Bestor, 2008). Although active DNA demethylation has been observed during early mouse embryogenesis, a DNA demethylase in human cells has not been identified (Reik et al., 2001). Studies in zebrafish and *Xenopus* models have produced promising indications that the DNA damage response gene Gadd45a may act as a DNA demethylase, although other reports dispute these findings (Barreto et al., 2007; Rai et al., 2008; Varambally et al., 2008).

In addition to altered gene expression or activity, global hypomethylation may also result from changes in the availability S-adenosylmethionine (SAM), a key substrate required for DNMT activity (Szyf et al., 2004). Methionine metabolism, an essential part of SAM biosynthesis, is often disrupted in transformed cells and may also be altered in this step-wise model of transformation (reviewed in Kim, 2005; Laird and Jaenisch, 1996).

An important factor distinguishing passive and active demethylation processes is the time period over which this loss of methylation takes place.

Active demethylation could take place in a short time frame in the absence of DNA replication as observed in early mouse embryogenesis (Oswald et al., 2000). Conversely, passive demethylation requires DNA replication, and would be expected to occur as a more gradual decline over time with each round of DNA replication. The mechanistic basis of hypomethylation in the MSC model may therefore be determined by investigating the time course over which methylation is lost following H-Ras^{V12} expression in MSC 4.

The mechanistic basis of genome-wide hypomethylation in MSC 5 may also be elucidated through analysis of changes in gene expression during step-wise transformation. Alterations in the expression of DNMTs, putative demethylases such as GADD45A or enzymes involved in SAM biosynthesis could be responsible for the hypomethylation observed in this system through the mechanisms described above. Microarray data have previously been obtained in all MSC cell lines, allowing further analysis of these data to be conducted in order to investigate epigenetic deregulation during transformation (Funes et al., 2007).

Although genome-wide hypomethylation in cancer has traditionally been viewed as an early event during transformation, few studies have definitively addressed whether hypomethylation is causative of transformation (Fearon and Vogelstein, 1990; Feinberg et al., 2006). Induction of genome-wide hypomethylation by *dnmt1* knockdown has been shown to increase tumour formation in mice, suggesting a causal relationship between the two events (Gaudet et al., 2003). However, conflicting evidence has been obtained regarding the occurrence of genome-wide hypomethylation in benign neoplasms, indicating

that this change is not always an early requirement of transformation (Gama-Sosa et al., 1983). Experiments with the conditional H-Ras^{V12} system will test whether transformation is dependent on genome-wide hypomethylation in the MSC model.

I hypothesize that repetitive element hypomethylation in MSC 5 results from a disrupted equilibrium of DNA methylation processes in these cells. This disruption is likely to result in a gradual loss of methylation following H-Ras^{V12} expression in MSC 4 cells but is not likely to be a requirement of transformation. I therefore aim to investigate the time course of hypomethylation following conditional and constitutive H-Ras^{V12} expression and assess whether hypomethylation is required for transformation to take place. Finally, I aim to investigate potential causes of genome-wide hypomethylation in this system through analysis of gene expression microarray data from all MSC cell lines.

4.2 Aims

1. To establish the time course over which hypomethylation takes place following H-Ras expression.
2. To determine whether cells can be transformed in the absence of genome-wide hypomethylation.
3. To conduct an analysis of gene expression microarray data in order to identify alterations to the expression of epigenetic regulators during step-wise transformation.

4.3 Results

4.3.1 Conditional H-Ras System – Cloning and Characterisation

A clear picture of the time course over which hypomethylation occurs in this model could provide an indication of the mechanistic basis behind this loss of methylation. If hypomethylation results from active mechanism (through direct removal of methyl groups from the DNA), loss of methylation could occur in a very short space of time. If, however, hypomethylation occurs in a 'passive' manner due to a loss of maintenance methylation, this loss would be expected to occur over a longer period of time through consecutive rounds of DNA replication.

The original MSC 5 population was created by infecting MSC 4 cells with retrovirus containing the oncogenic H-Ras^{V12} construct and selecting successfully infected cells. The time period over which infection and selection takes place makes it difficult to establish exactly when selected cells begin to express H-Ras^{V12}, or the point at which genome-wide hypomethylation takes place. A more efficient way to explore the time course of demethylation following H-Ras expression would be to use a conditional system. To this end, a conditional H-Ras^{V12} construct was obtained from Dr. Julian Downward (CR-UK London Research Institute). This construct consisted of the full coding sequence for oncogenic H-Ras^{V12} and a mutated form of the oestrogen receptor (ER^{TM}) expressed together as a fusion protein. ER^{TM} is sensitive to 4-hydroxytamoxifen (4-OHT) but not to the natural ligand for the wild type receptor (17β-estradiol) (Danielian et al., 1993). In the absence of 4-OHT, the fusion protein is

complexed with hsp90 in the cytoplasm, preventing downstream H-Ras signaling. This interaction is abolished when treated with hydroxytamoxifen, allowing active H-Ras signaling to take place (Mattioni et al., 1994).

In order to use this construct in the MSC model, the ER^{TM}-$H\text{-}Ras^{V12}$ fragment (ER-Ras) had to be subcloned from the original vector (pLZRS) into the pWZL retroviral vector (containing a hygromycin resistance cassette) previously used for H-Ras overexpression in MSC 4. Sequencing data previously obtained from the pLZRS vector indicated that the ER^{TM}-H-Ras fragment was flanked by BamHI and SalI restriction enzyme sites (David Hancock, personal communication), which were also present in the multiple cloning site of pWZL. Double digestion of pLZRS-ER-Ras released a fragment of the expected size (1.6kb), which was gel extracted and ligated into pWZL previously linearised with BamHI and SalI (Figure 4.1a). Following transformation of bacteria with the ligation product, individual colonies were picked and maxipreps of the plasmid were prepared. Diagnostic digestion of the final plasmids indicated that a fragment of the correct size had been cloned into pWZL, with the identity of this fragment confirmed by sequencing (Figure 4.2b). The sequencing data also showed that no mutations were introduced to the sequence during cloning, and the entire sequence was in frame (data not shown).

Retroviruses containing the pWZL-ER^{TM}-$H\text{-}Ras^{V12}$ construct and empty vector control (pWZL) were produced and used to infect MSC 4 cells. Following selection of infected cells with hygromycin, extensive characterisation of ER^{TM}-$H\text{-}Ras^{V12}$ activity was conducted. To test that expression of the fusion protein could be induced by 4-OHT treatment, MSC 4 containing ER^{TM}-$H\text{-}Ras^{V12}$ (MSC $4^{ER\text{-}Ras}$) or empty vector (MSC 4^{pWZL}) were treated with 100nM 4-OHT for 24

hours before total protein was harvested and used for western blot analysis. When the membrane was probed with an anti-H-Ras antibody, a band of the expected size (50kDa) was present in MSC $4^{ER\text{-}Ras}$ cells treated with 4-OHT but not in empty vector control cells. A small basal level of ER^{TM}-H-Ras expression was evident in untreated MSC $4^{ER\text{-}Ras}$ cells (Figure 4.1c). To test whether the expression of this fusion protein resulted in functional downstream MAPK signaling, levels of phospho-ERK1/2 (pERK) were assessed by western blot and compared to total ERK1/2 levels. This indicated that MSC $4^{ER\text{-}Ras}$ cells treated with 4-OHT contained higher levels of pERK compared to total ERK, demonstrating elevated MAPK signaling relative to untreated cells and empty vector control cells (Figure 4.1c).

In order to establish the optimal concentration of 4-OHT to use with these cell lines, ER^{TM}-H-Ras, pERK and total ERK levels were assessed 24 hours after incubation with increasing concentrations of 4-OHT (Figure 4.2a). Elevated expression of the fusion protein was observed in MSC $4^{ER\text{-}Ras}$ cells treated with increasing concentrations of 4-OHT, with levels of pERK relative to total ERK increasing in parallel. No expression of the fusion protein or increase in pERK levels was observed in MSC 4^{pWZL} cells treated with the same concentrations of 4-OHT (Figure 4.2b).

Similarly, ER^{TM}-H-Ras and pERK levels were assessed over time following the addition of 4-OHT. Analysis of protein levels in MSC $4^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells treated with 100nM 4-OHT for 72 hours revealed increasing levels of ER^{TM}-H-Ras over time, with this increase first occurring 12 hours after 4-OHT addition and reaching maximal levels after 48 hours (Figure 4.2c). No

fusion protein was detectable in MSC 4^{pWZL} cells treated with 4-OHT for the same time (Figure 4.2d).

The effect of ER^{TM}-H-Ras expression on cell viability was assessed by conducting a MTS assay on MSC $4^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells treated with 100nM and 500nM 4-OHT or equivalent amounts of ethanol over the course of 72 hours (Figure 4.3 a & b). In all cases, these results show no significant difference in cell growth between conditions, indicating that 4-OHT treatment and ER^{TM}-H-Ras expression do not adversely affect cell viability.

Finally, an *in vitro* assay of transformation was conducted by assessing the growth of cells in soft agarose. MSC $4^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells were seeded in soft agarose in the presence or absence of 100nM 4-OHT, with the number of colonies quantified after 12 days. This experiment demonstrated that MSC $4^{ER\text{-}Ras}$ cells grown in the presence of 100nM 4-OHT were able to form colonies in soft agarose, whereas untreated and vector control cells were not (Figure 4.3 c & d). Colony formation by MSC $4^{ER\text{-}Ras}$ cells was also assessed at different concentrations of 4-OHT, indicating a modest increase in colony number on increasing concentrations of 4-OHT (Figure 4.3e). This demonstrates that MSC $4^{ER\text{-}Ras}$ cells are transformed when H-Ras signaling is induced by 4-OHT treatment.

Figure 4.1

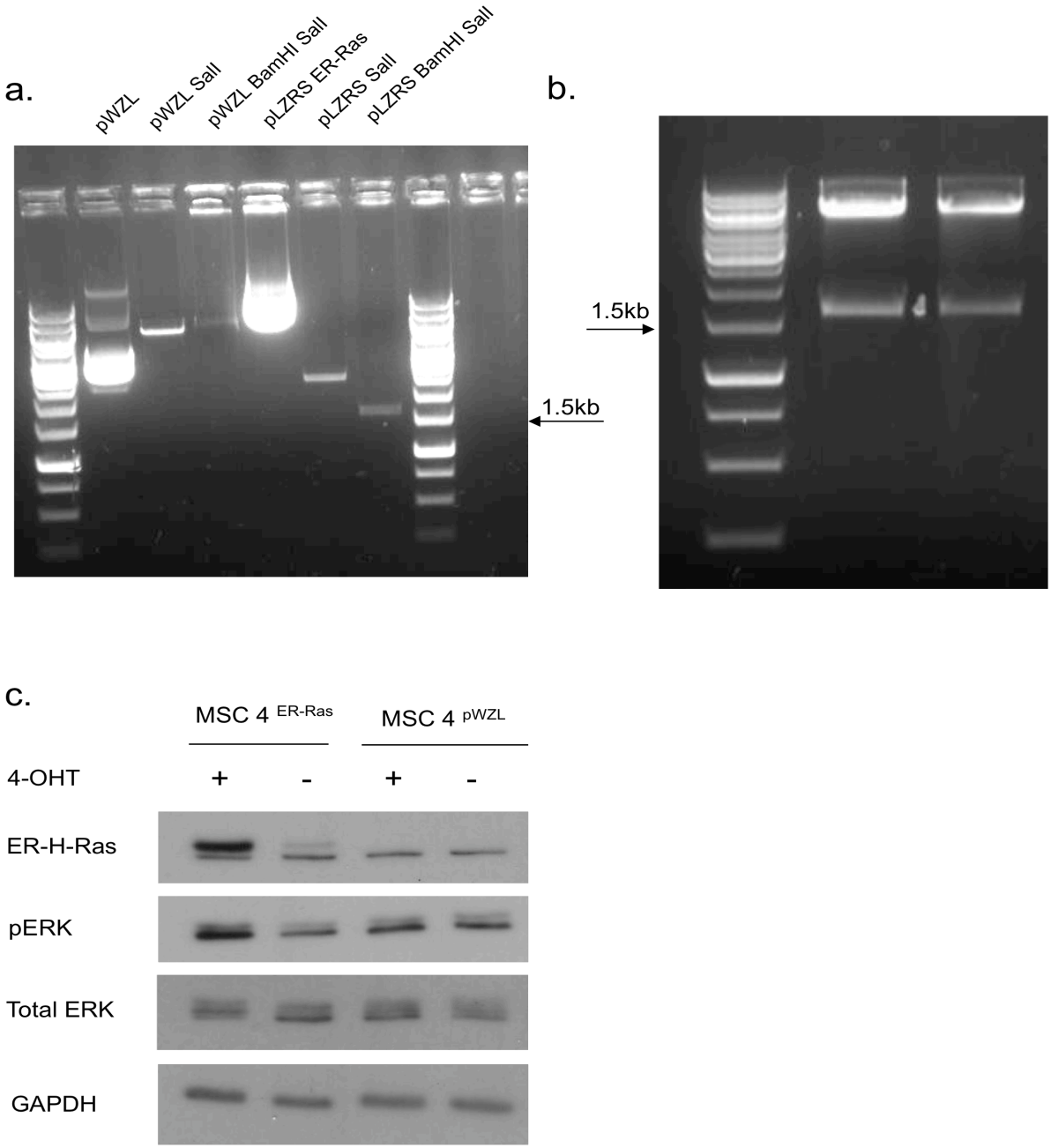


Figure 4.1: Cloning of ER^{TM}-H-Ras fragment from pLZRS to pWZL and characterisation of ER^{TM}-H-Ras activity. (a) Agarose gel electrophoresis of restriction enzyme digestion products. The ER^{TM}-H-Ras fragment (1.6kb in length) was removed from pLZRS by restriction enzyme digestion with BamHI and SalI. The target vector (pWZL) was linearised by BamHI and SalI digestion before the ER^{TM}-H-Ras fragment was ligaed into the vector backbone. **(b)** Successful ligation products were confirmed by diagnostic digest with BamHI and SalI. **(c)** Western blot analysis of MSC 4 cells infected with retrovirus containing ER^{TM}-H-Ras (MSC 4$^{ER\text{-}Ras}$) or empty vector (MSC 4^{pWZL}), treated with 4-OHT (+) or ethanol (-) for 24 hours. Levels of H-Ras, pERK, total ERK and GAPDH (loading control) were assessed using antibodies against each of these proteins.

Figure 4.2

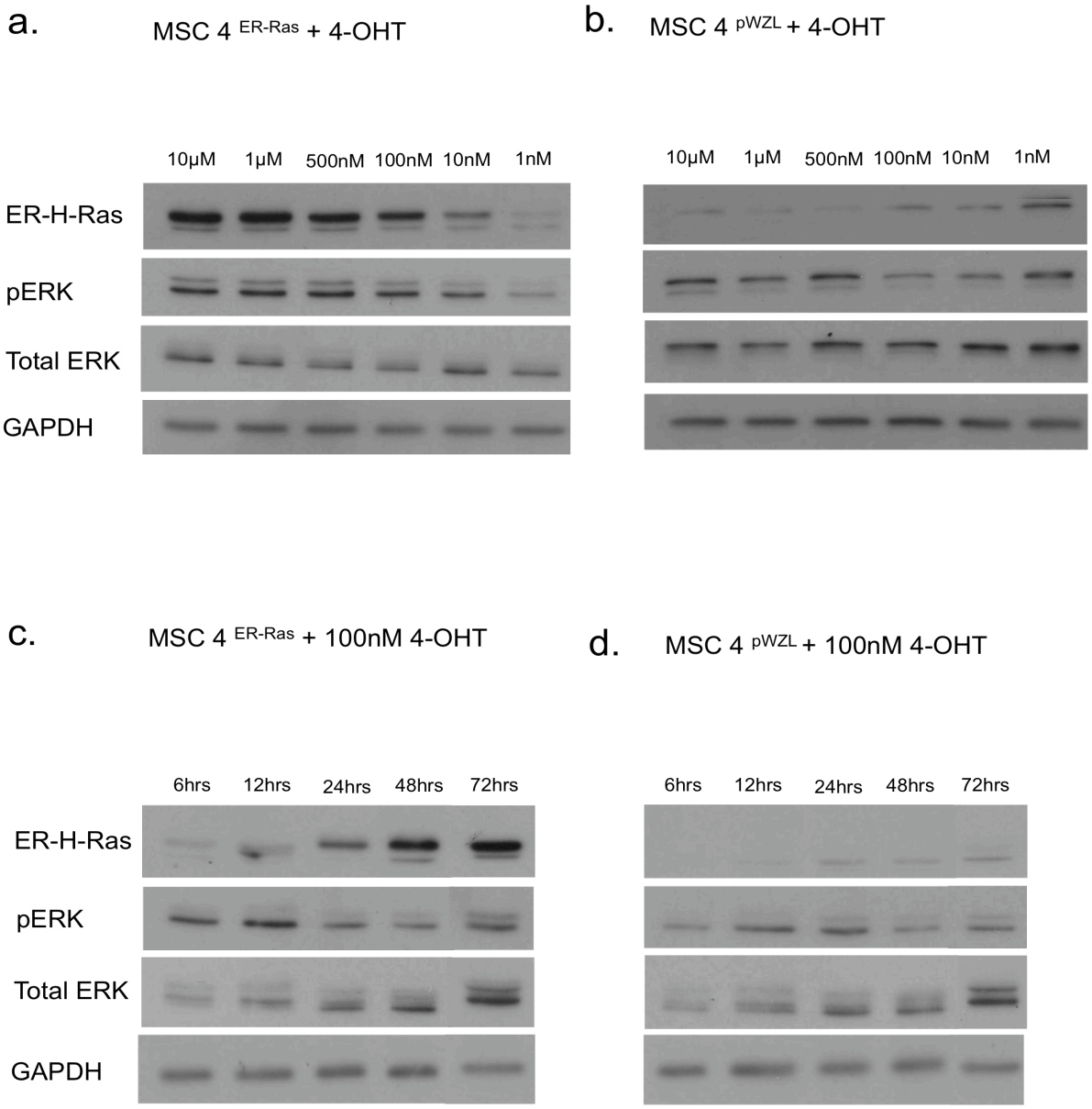


Figure 4.2: Characterisation of conditional ER^{TM}-H-Ras dose and time response. (a-b) Western blot analysis of MSC 4$^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells treated with increasing concentrations of 4-OHT for 24 hours. Levels of H-Ras, pERK, total ERK and GAPDH (loading control) were assessed using antibodies against each of these proteins. **(c-d)** Western blot analysis of MSC 4$^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells were treated with 100nM 4-OHT over the course of 72 hours. Levels of H-Ras, pERK, total ERK and GAPDH (loading control) were assessed using antibodies against each of these proteins.

Figure 4.3

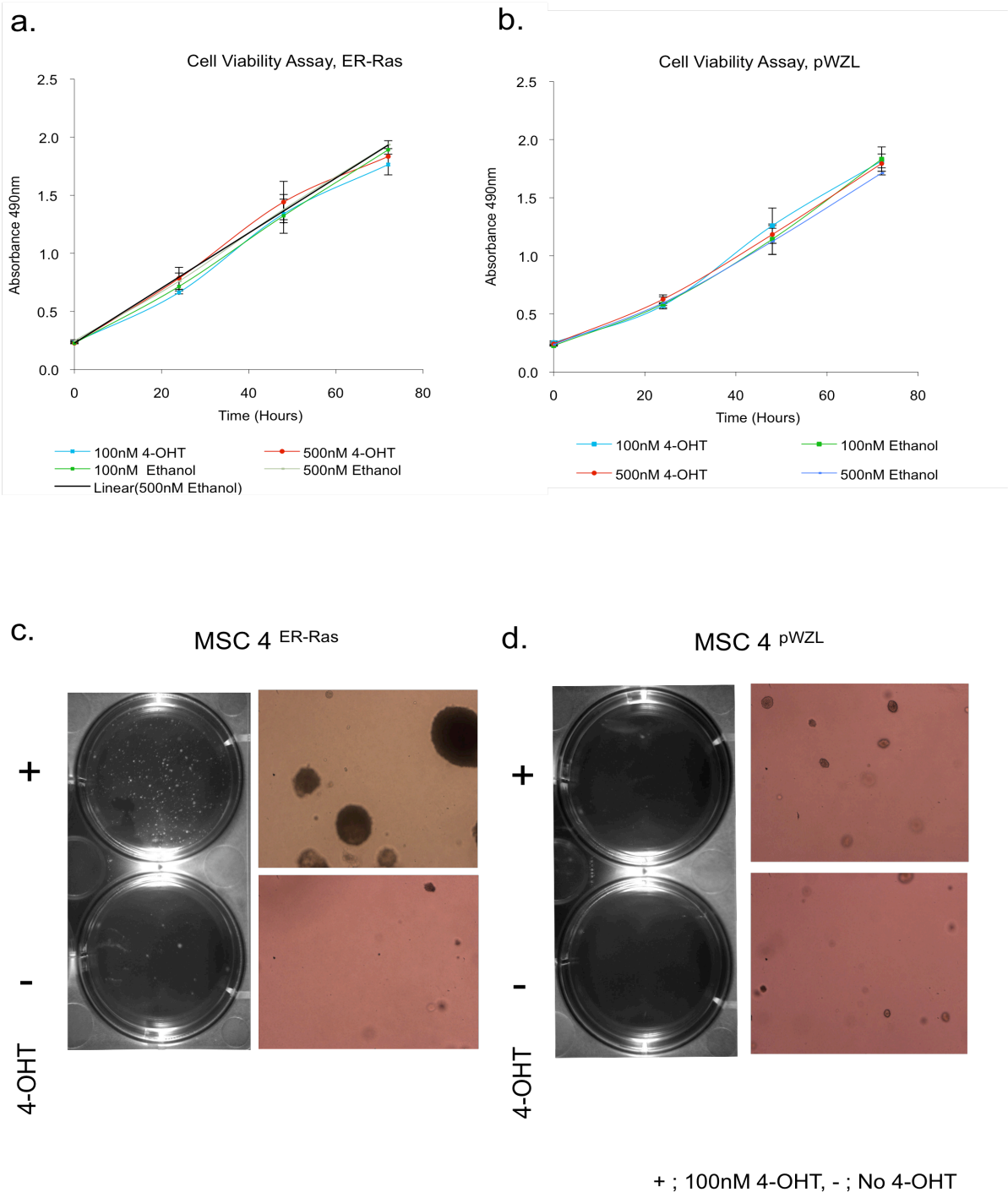


Figure continued on the following page

Figure 4.3

e.

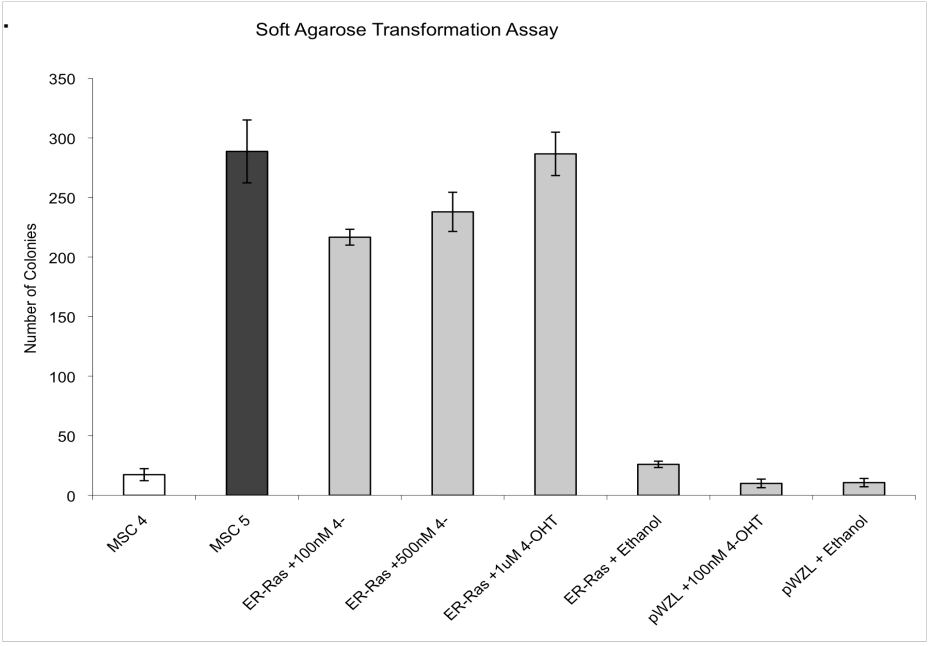


Figure 4.3: Cell viability and growth of MSC $4^{ER\text{-}Ras}$ and MSC 4^{pWZL} is not affected by 4-OHT and MSC $4^{ER\text{-}Ras}$ treated with 4-OHT are transformed. Cell viability (MTS) assay for **(a)** MSC $4^{ER\text{-}Ras}$ and **(b)** MSC 4^{pWZL} seeded at a density of 1000 cells per well and grown over the course of 72 hours in the presence of 100 and 500nM 4-OHT or equivalent volumes of ethanol. Each point represents the average absorbance (490nm) from 5 replicates when background absorbance is subtracted. **(c)** Soft agarose transformation assay of MSC $4^{ER\text{-}Ras}$ and **(d)** MSC 4^{pWZL} cells seeded at a density of 10,000 cells per well and grown for 12 days in the presence of 100nM 4-OHT (+) or an equivalent volume of ethanol (-). Positive control (MSC 5) and negative control (MSC 4) cells were grown in parallel (data not shown). Whole well photographs are depicted alongside 40x magnification. **(e)** Soft agarose transformation assay of MSC $4^{ER\text{-}Ras}$ and MSC 4^{pWZL} repeated as above, but with increasing concentrations of 4-OHT. The number of colonies after 12 days growth at 37°C was quantified using the SynGene GeneTools image analysis software. Columns depict average of three replicates; Error bars depict SEM.

4.3.2 Global Methylation Levels Following ER^{TM}-H-Ras Induction

To investigate whether repetitive element hypomethylation occurred following ER^{TM}-H-Ras induction and to establish the time course of this change, MSC 4$^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells were grown in the presence and absence of 500nM 4-OHT for one month, with methylation of Alu, LINE1 and Sat2 elements assayed by MethyLight 72 hours after the addition of 4-OHT and every week for four weeks thereafter (Figure 4.4a-c). These results indicated that Alu, LINE1 and Sat2 methylation did not differ significantly between MSC 4$^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells treated with 4-OHT at any of the time points. This result was confirmed by bisulphite sequencing of the Sat2 element (Figure 4.4d). Continued expression of the fusion protein and elevated pERK levels in MSC 4$^{ER\text{-}Ras}$ cells were confirmed at the final time point by western blot (Figure 4.4e).

Figure 4.4

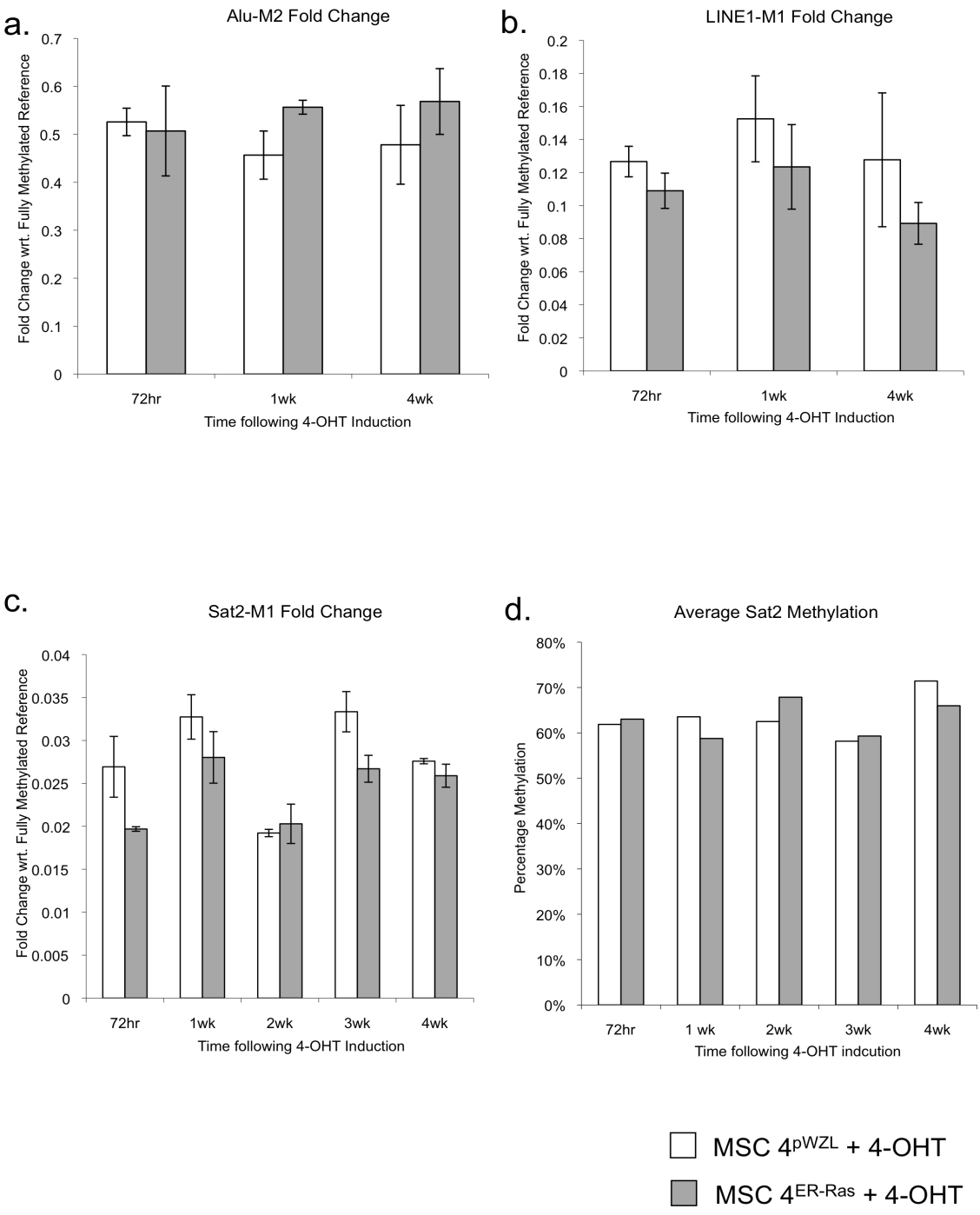


Figure continued on the following page

Figure 4.4

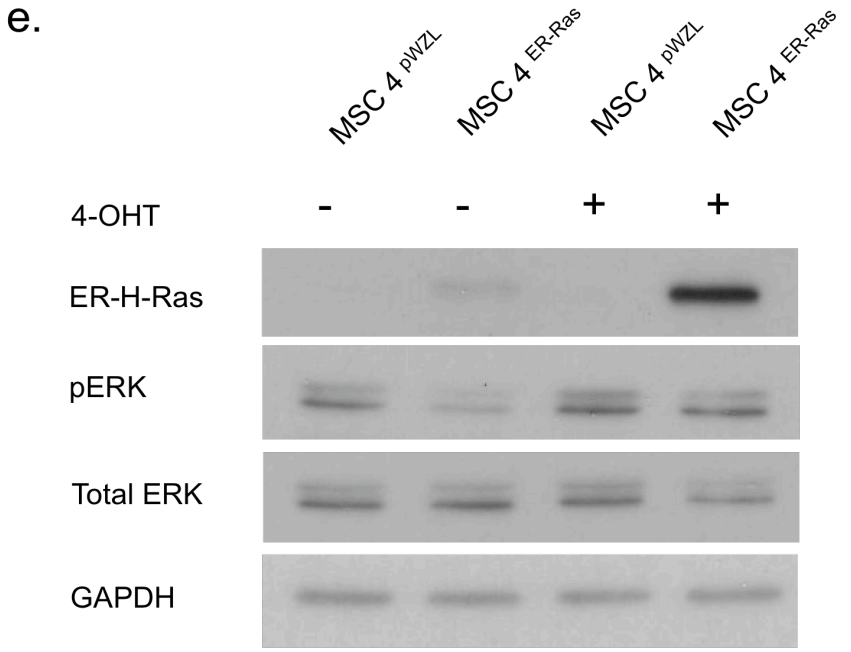


Figure 4.4: Sat2, LINE1 and Alu methylation does not change following induced transformation of MSC $4^{ER\text{-}Ras}$ cells. MSC $4^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells were grown in the presence of 500nM 4-OHT for one month, with DNA harvested for methylation analysis at various time points during this period. MethyLight was conducted using primers against methylated **(a)** Alu **(b)** LINE1 and **(c)** Sat2 elements in bisulphite converted DNA from MSC $4^{ER\text{-}Ras}$ and MSC 4^{pWZL} at each time point, alongside fully unmethylated DNA as a negative control (not shown). Fold-change in methylation was calculated in triplicate by ΔΔCt using an unbiased Alu reaction as a loading control and normalised to fully methylated (*M.SssI* treated) reference DNA. Student's T-test was performed for all statistical analyses. Columns display average of three replicates; Error bars display SEM. **(d)** Bisulphite sequencing confirmation of the MethyLight data. Sat2 elements were amplified by PCR from bisulphite modified genomic DNA with individual PCR products cloned into pGEM-T and sequenced. Average methylation is expressed as a percentage of all CG sites assayed by sequencing 12 PCR products. Columns display average percentage methylation; Statistical analysis was conducted using the Paired Wilcoxon Signed Rank Test. **(e)** Western blot analysis of ER-H-Ras, pERK, total ERK and GAPDH (loading control) protein levels in all cell lines at the final time point (four weeks following 4-OHT induction).

4.3.3 *H-Ras^{V12} Overexpression in MSC 4 and Sat2 Methylation Analysis*

To determine whether the repetitive element hypomethylation observed in the original MSC 5 cells could be recapitulated, oncogenic H-Ras^{V12} was overexpressed in MSC 4 cells. Retrovirus containing H-Ras^{V12} within the pWZL vector was generated alongside empty vector control virus. Infection of MSC 4 cells was performed in triplicate for each virus, with hygromycin resistant cells selected over the course of one week following infection. After selection, H-Ras^{V12} overexpression was confirmed by western blot and transformation of the cells was assessed by anchorage independent growth in soft agarose (Figure 4.5a-c). This indicated that MSC 4 cells expressing H-Ras^{V12} were transformed whereas cells containing the empty vector control were not.

Sat2 methylation levels in these cells were assayed by MethyLight, indicating no significant difference between H-Ras^{V12} and empty vector containing cells one week following infection (Figure 4.6a). To determine whether Sat2 hypomethylation would take place over a longer period of time, all cells were cultured as normal for four weeks, with Sat2 methylation levels assessed at this time point by MethyLight (Figure 4.6b). These data demonstrated a significant reduction in Sat2 methylation between MSC 4 cells expressing H-Ras^{V12} and empty vector control cells.

Figure 4.5

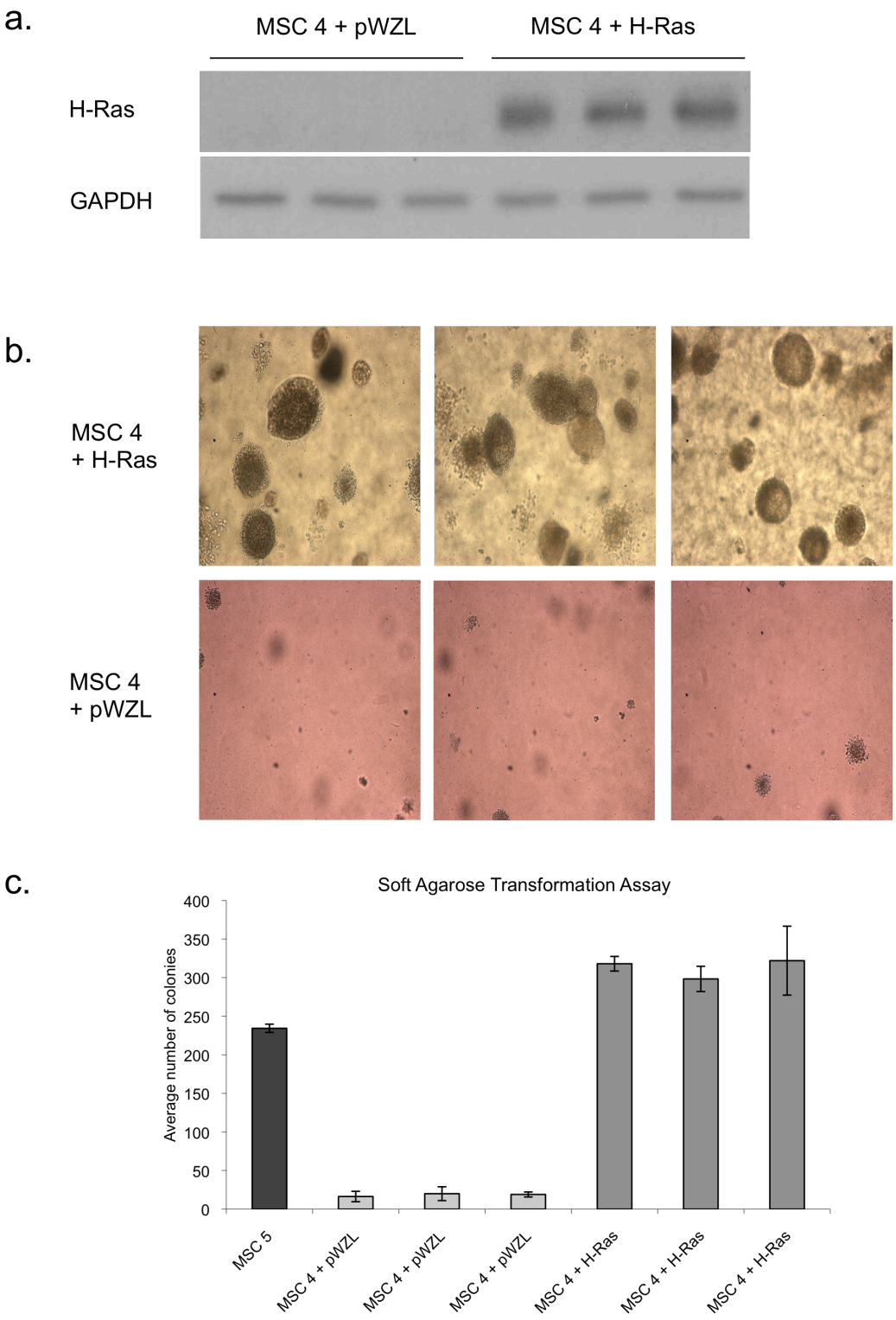


Figure 4.5: Overexpression of H-Ras^{V12} in MSC 4 cells induces transformation. Retrovirus containing H-Ras^{V12} within the pWZL vector was generated and used to infect MSC 4 cells in triplicate alongside empty vector control virus. **(a)** Western blot analysis of H-Ras and GAPDH

(loading control) protein levels following infection and selection of MSC 4 cells. **(b)** Soft agarose transformation assay of MSC 4 cells expressing H-Ras^{V12} (MSC 4 + H-Ras) or empty vector control (MSC 4 + pWZL) seeded at a density of 10,000 cells per well and grown for 12 days at 37°C. Images at 40x magnification. **(c)** Quantification of colony number from soft agarose assay using SynGene GeneTools image analysis software.

Figure 4.6

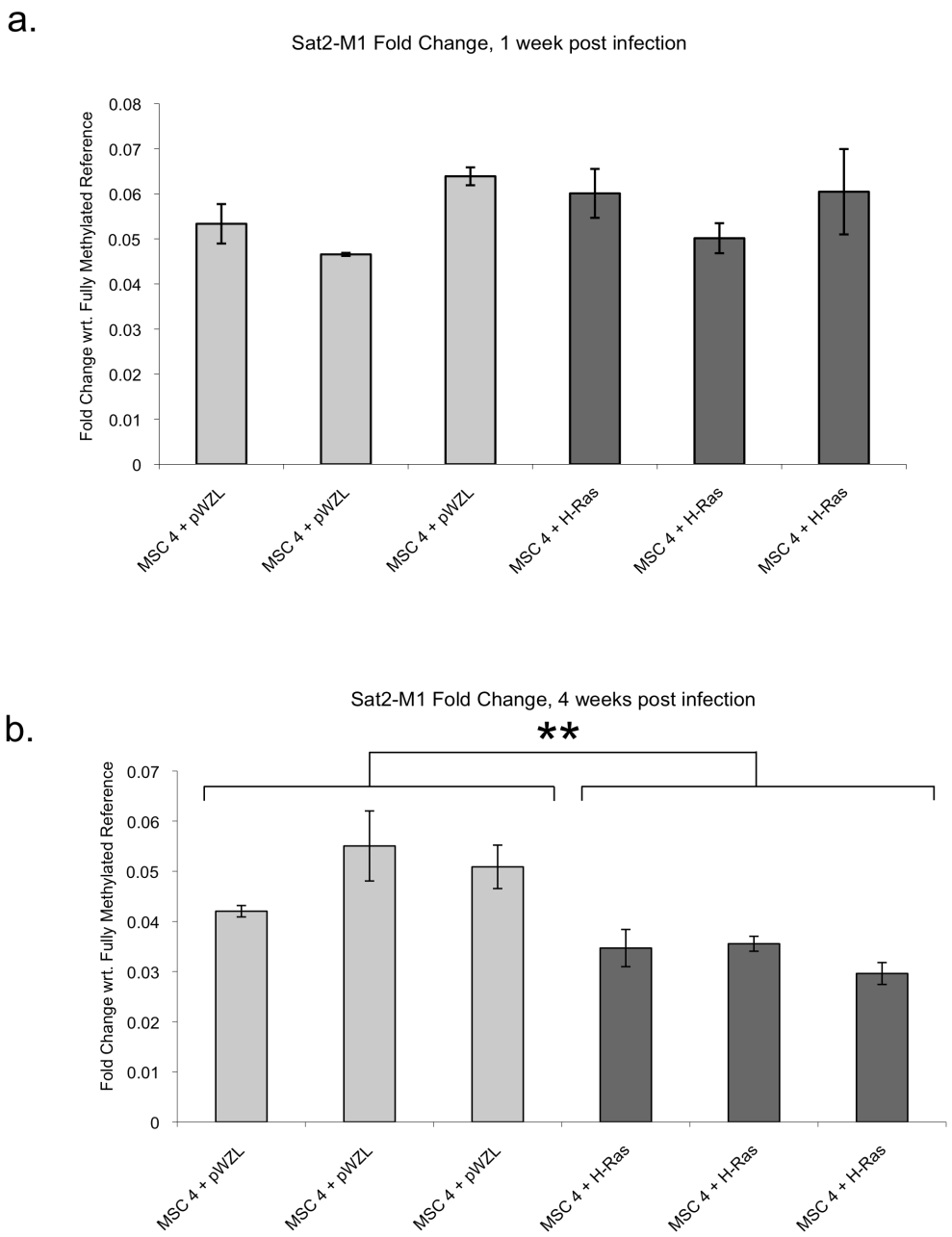


Figure 4.6: Sat2 hypomethylation is induced following H-Ras^{V12} expression in MSC 4 cells and continued culturing for four weeks. MethyLight was conducted using primers against methylated Sat2 elements in bisulphite converted DNA from MSC 4 cells expressing H-Ras^{V12} (MSC 4 + H-Ras) or empty vector control (MSC 4 + pWZL) **(a)** one week and **(b)** four weeks following retroviral infection, alongside fully unmethylated DNA as a negative control (not shown). Fold-change in methylation was calculated in triplicate by ΔΔCt using an unbiased Alu reaction as a loading control and normalised to fully methylated (M.SssI treated) reference DNA. Student's T-test was performed for all statistical analyses. Columns display average of three replicates; Error bars display SEM.

4.3.4 Analysis of gene expression microarray data

In order to establish whether changes in the expression of epigenetic regulators during step-wise transformation were responsible for repetitive element methylation in MSC 5, analysis of gene expression microarray data for all MSC cell lines was conducted.

Gene expression microarray data were previously obtained for all MSC lines in triplicate on Affymetrix U133+2 arrays (Funes et al., 2007). Significant changes in gene expression between MSC 0 and MSC 5 were determined by FDR corrected T-Test of log expression values between conditions, with $q < 0.01$ considered significant. In order to focus the analysis on relevant genes involved in epigenetic regulation, a list of epigenetic regulators was compiled from an extensive literature review, comprising 967 probe sets for 341 genes (Table A2). Expression data for these probe sets were extracted from the genome-wide list of differentially expressed genes.

I identified 106 probe sets, representing 69 genes, which were significantly up- or down- regulated between MSC 0 and MSC 5 (Figure 4.7 & Table A3). Of the DNA methyltransferase genes, only *DNMT1* showed a significant change in expression between MSC 0 and MSC 5, with a significant increase in expression (1.7 fold increase, $q = 1x10^{-6}$). qRT-PCR analysis confirmed this result, demonstrating that *DNMT1* mRNA expression was up-regulated in MSC 5 compared to MSC 0, increasing significantly at MSC 2E7 on pRb inhibition, with this increase maintained through subsequent oncogenic hits (Figure 4.8a). In comparison, the *de novo* methyltransferases DNMT3a, 3b and 3l did not display any significant changes in expression between MSC 0 and

MSC 5. Reduced DNMT1 expression is therefore unlikely to account for the hypomethylation observed in this model.

Analysis of gene expression microarray data also indicated a significant up-regulation of *GADD45A*, a putative DNA demethylase, between MSC 0 and MSC 5 (2.6 fold increase, q = 0; Figure 4.7a). This increase occurred between MSC 4 and MSC 5, the same point at which repetitive element hypomethylation was observed. Analysis of *GADD45A* mRNA levels by qRT-PCR confirmed this result, demonstrating a significant up-regulation following H-Ras^{V12} expression (Figure 4.8b). Furthermore, qRT-PCR data obtained from the conditional H-Ras model also demonstrated a significant up-regulation of *GADD45A* expression one week following 4-OHT induction (Figure 4.8c).

Analysis of gene expression microarray data for enzymes involved in the metabolic pathways linked to methionine production also indicated significant changes to the expression of these genes (Figure 4.9). A number of genes coding for enzymes involved in the conversion of metabolites away from production of methionine and towards *de novo* purine biosynthesis are up-regulated in MSC 5 compared to MSC 0. Additionally, *MTHFR*, a gene involved in the conversion of 5,10-methylenetetrahydrofolate to the methyl donor 5-methyltetrahydrofolate is significantly down-regulated in MSC 5 relative to MSC 0.

Figure 4.7

a.

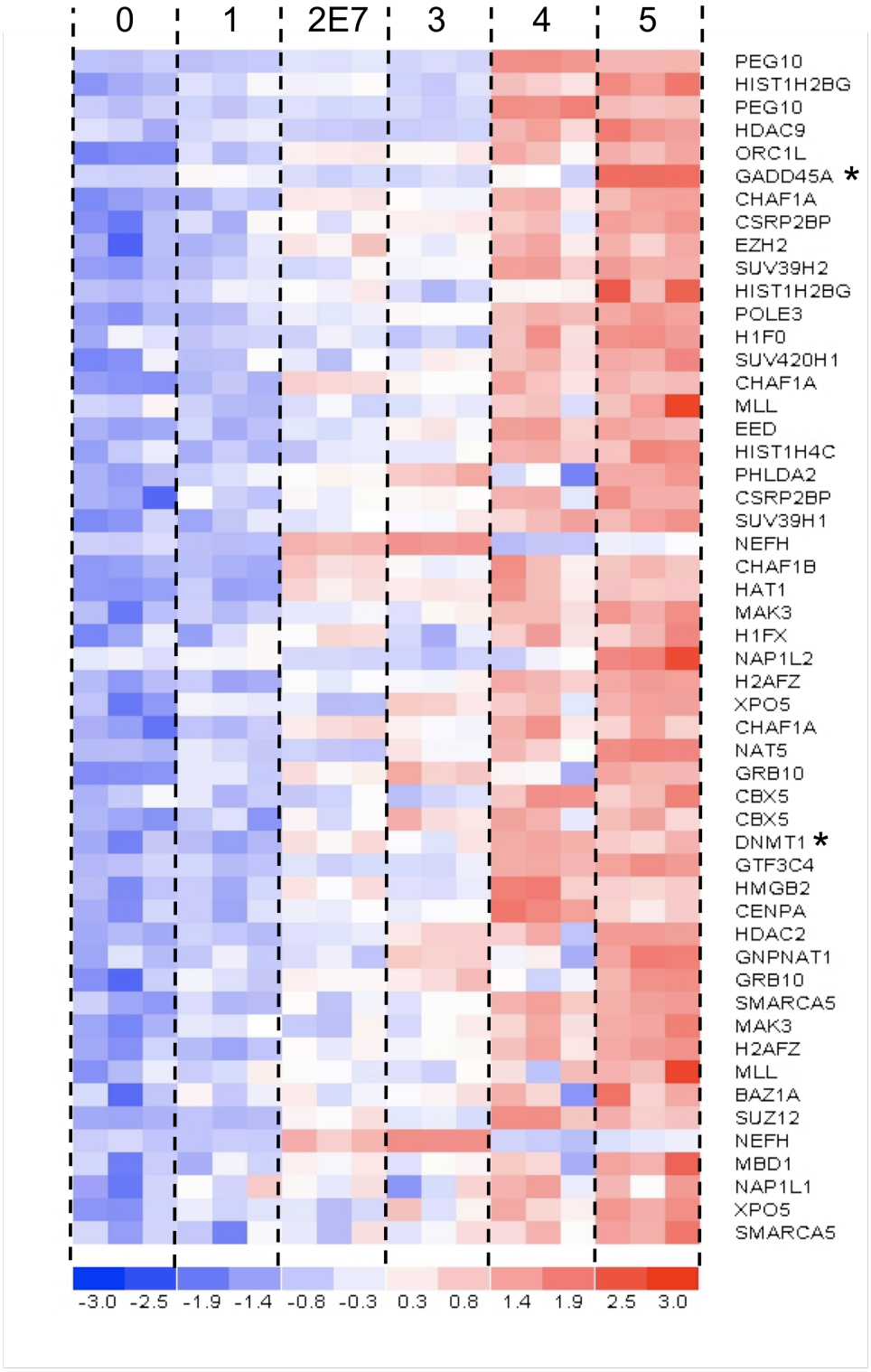

0
1
2E7
3
4
5
PEG10
HIST1H2BG
PEG10
HDAC9
ORC1L
GADD45A *
CHAF1A
CSRP2BP
EZH2
SUV39H2
HIST1H2BG
POLE3
H1F0
SUV420H1
CHAF1A
MLL
EED
HIST1H4C
PHLDA2
CSRP2BP
SUV39H1
NEFH
CHAF1B
HAT1
MAK3
H1FX
NAP1L2
H2AFZ
XPO5
CHAF1A
NAT5
GRB10
CBX5
CBX5
DNMT1 *
GTF3C4
HMGB2
CENPA
HDAC2
GNPNAT1
GRB10
SMARCA5
MAK3
H2AFZ
MLL
BAZ1A
SUZ12
NEFH
MBD1
NAP1L1
XPO5
SMARCA5
-3.0 -2.5 -1.9 -1.4 -0.8 -0.3 0.3 0.8 1.4 1.9 2.5 3.0

Figure continued on the following page

Figure 4.7

b.

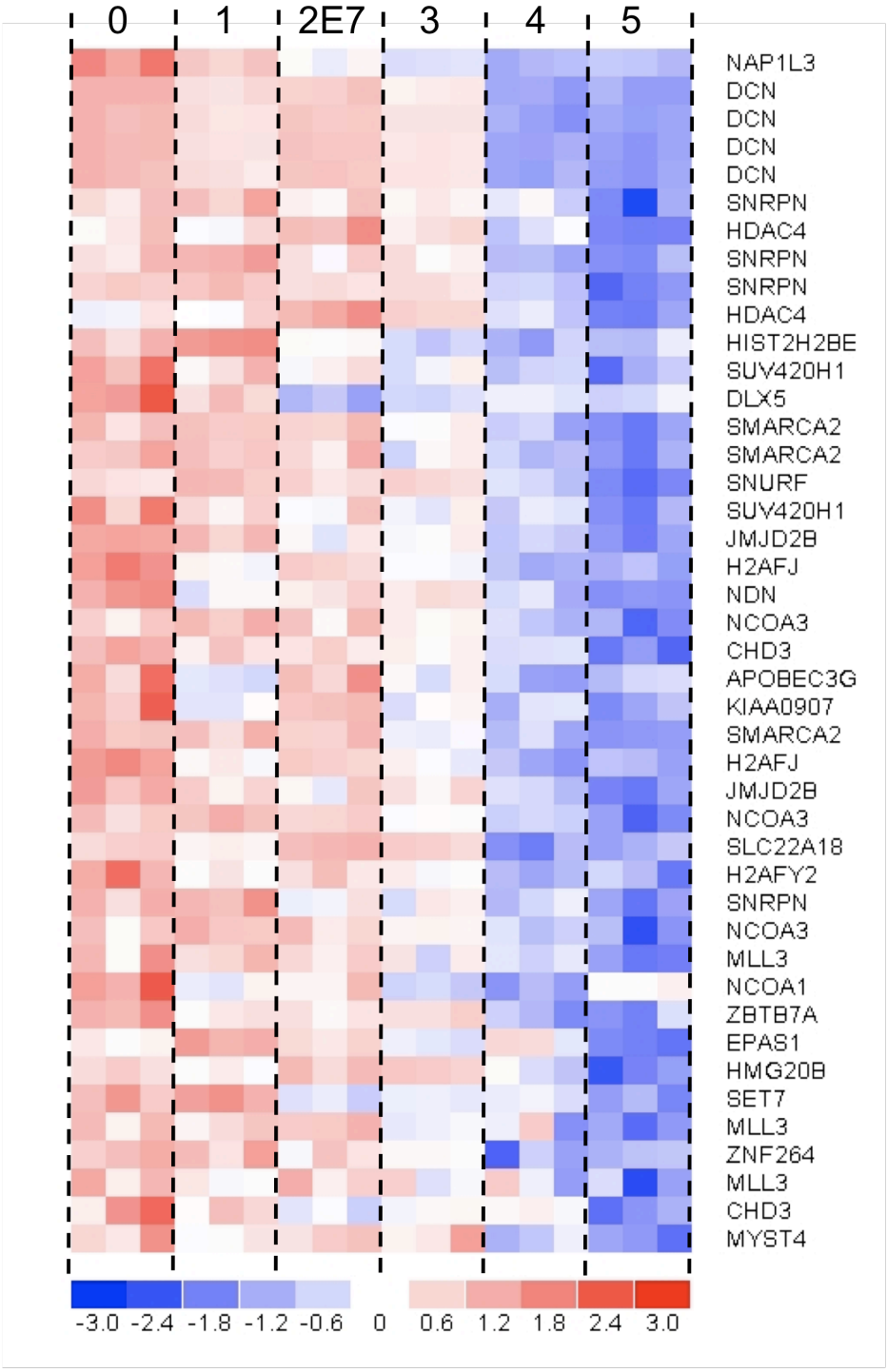


Figure 4.7: Expression of epigenetic regulator genes changes during step-wise transformation. Gene expression microarray data for epigenetic regulators was analysed for significant changes in expression between MSC 0 and MSC 5. Heatmaps showing triplicate gene expression microarray data for significantly **(a)** up-regulated and **(b)** down-regulated genes in each of the five transforming steps (MSC 0 - 5).

Figure 4.8

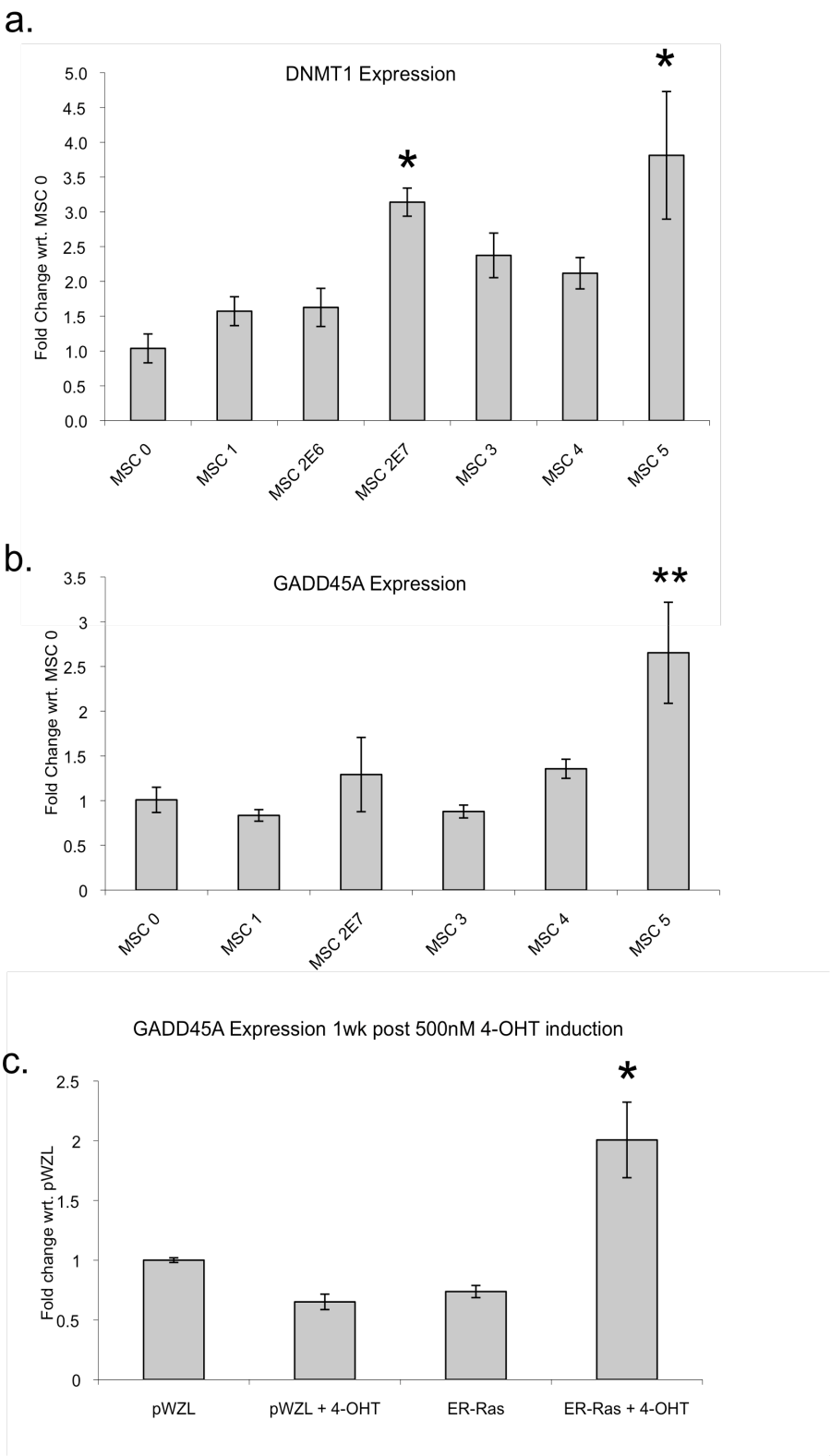


Figure 4.8: DNMT1 and GADD45a expression increase significantly during step-wise transformation of MSC. (a) Quantitative RT-PCR confirmation of DNMT1 expression. Fold-change in DNMT1 expression with respect to MSC 0 was calculated in triplicate by ΔΔCt using GAPDH as a loading control and normalised to parental MSC. **(b)** Fold-change in GADD45A

expression with respect to MSC 0 was calculated in triplicate by ΔΔCt using GAPDH as a loading control and normalised to parental MSC. **(c)** MSC $4^{ER\text{-}Ras}$ and MSC 4^{pWZL} cells grown in the presence of 500nM 4-OHT or an equivalent volume of ethanol for one week. GADD45A expression was assessed by qRT-PCR as described above, with fold change calculated relative to MSC 4^{pWZL} cells. Student's T-test was performed for all statistical analyses. * $p < 0.05$ with respect to MSC 0. Columns display average of three replicates; Error bars display SEM.

Figure 4.9

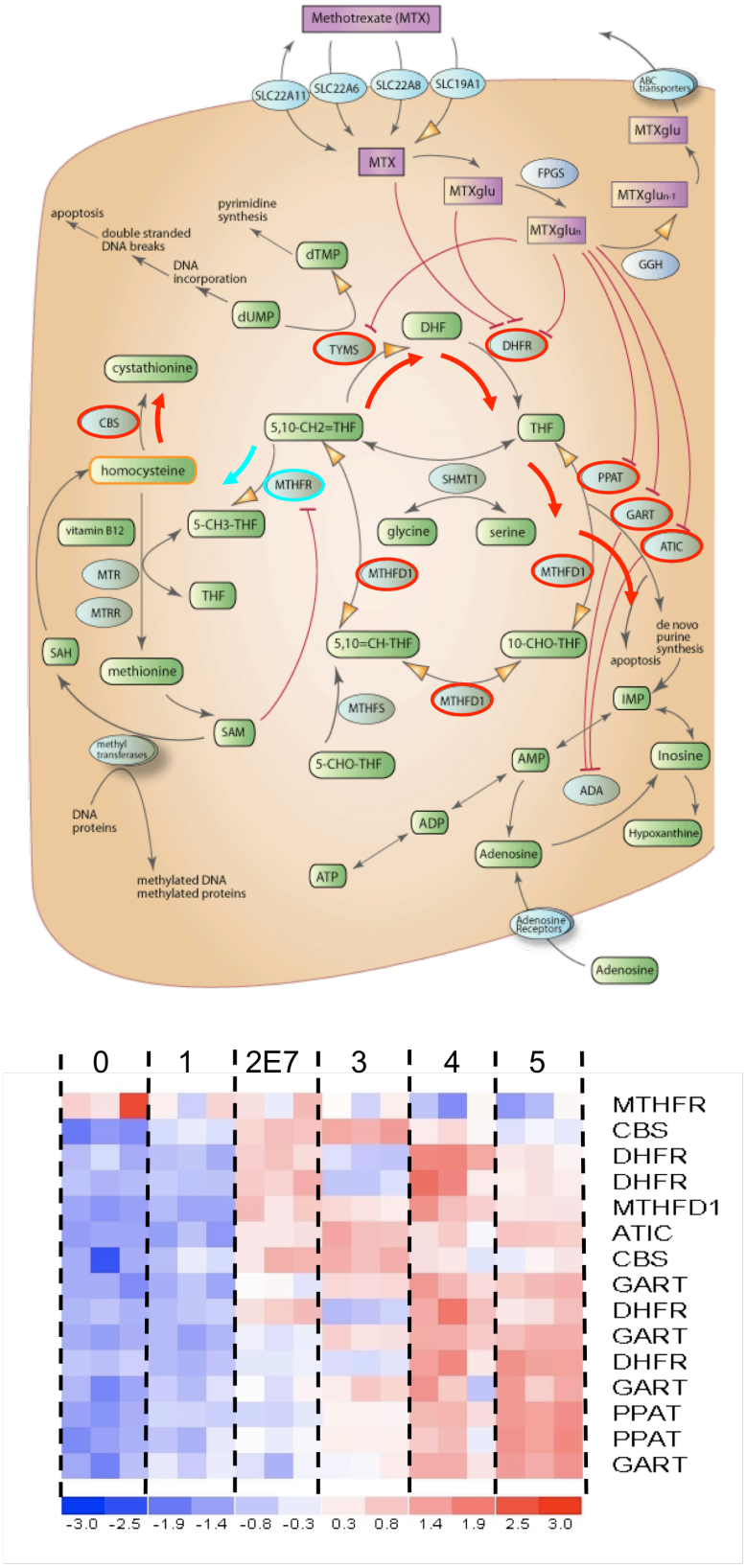


Figure 4.9: Enzymes involved in the methotrexate pathway show a distinct pattern of regulation towards *de novo* purine synthesis and away from S-adenosylmethionine (SAM) synthesis. (a) Schematic showing the substrates and enzymes of biochemical reactions

surrounding the methotrexate pathway and SAM biosynthesis. Genes which show a significant increase in expression between parental and transformed MSC are highlighted in red, with genes displaying reduced expression highlighted in blue. Figure adapted from PharmGKB website (www.pharmgkb.org) and *Flanagan et al.* 2009. **(b)** Heatmaps showing triplicate gene expression microarray data for significantly down-regulated and up-regulated genes in each of the five transforming steps (MSC 0 - 5).

4.3.5 Correction of GEM Data for Cell Proliferation Rate

In order to account for differences in proliferation rate between MSC cell lines and changes in gene expression resulting from this difference, correction of the GEM data was conducted. Log expression values for all probe sets were adjusted based on expression values of the proliferation-linked marker PCNA in each cell line. The adjusted array data were subjected to statistical analysis as described above to extract significant changes in epigenetic regulators between MSC 0 and MSC 5 (Figure 4.10).

This analysis indicated significant changes in the expression of 383 probe sets, representing 203 genes, with the majority of genes down-regulated over the course of step-wise transformation (Figure 4.10a & Table A4) and 15 genes up-regulated (Figure 4.10b). Furthermore, there is a significant difference between the genes identified in this analysis compared to analysis of the unadjusted expression data. For example, expression of *DNMT1* does not significantly change when corrected for proliferation ($q = 0.17$), whereas expression of *DNMT3a* decreases between MSC 0 and MSC 5 ($q = 0.002$). Some of the changes identified in the previous analysis, such as increasing expression of *GADD45a* between MSC 4 and MSC 5 ($q = 0$), still achieve statistical significance when expression values are corrected for proliferation.

Figure 4.10

a.

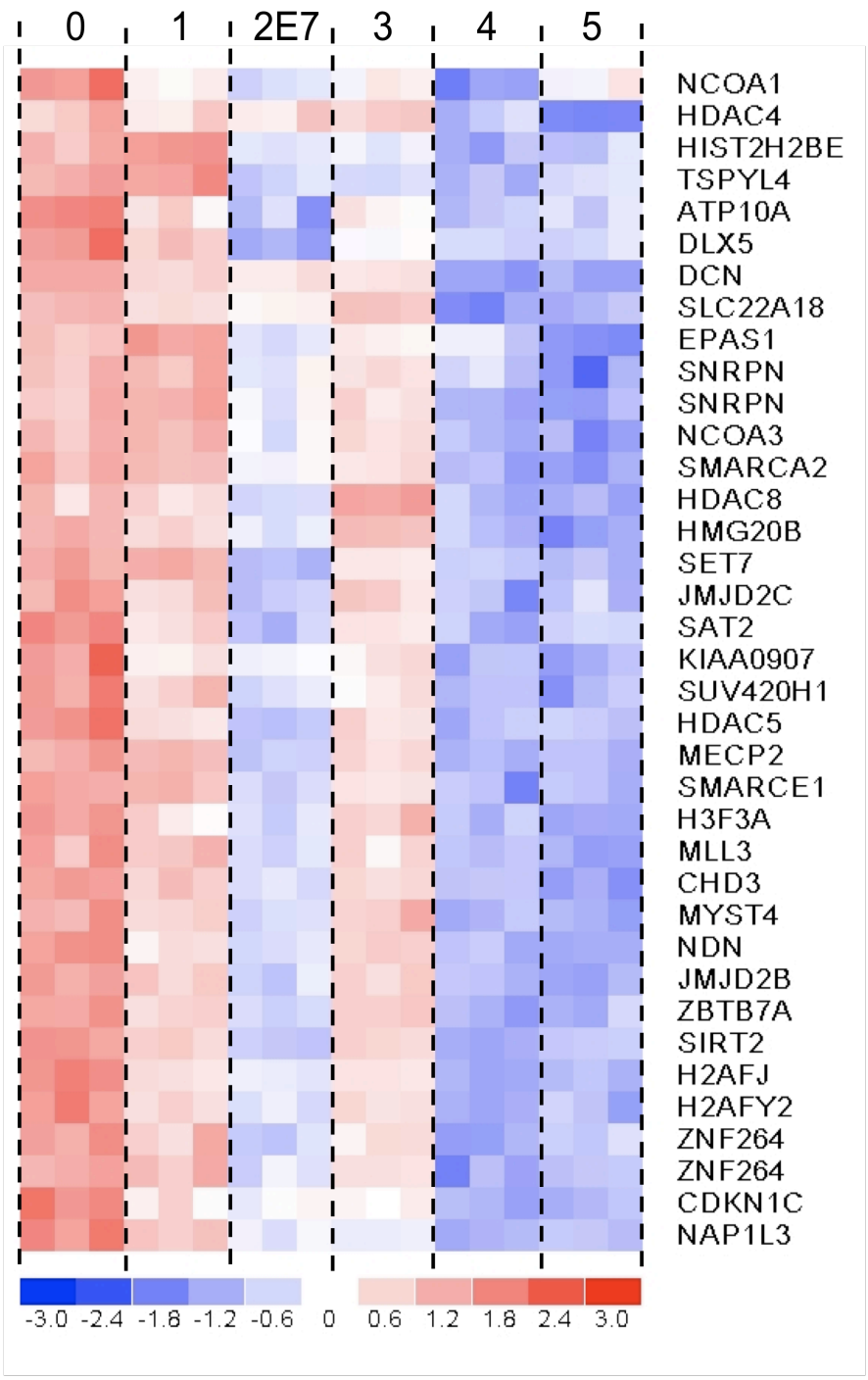

0
1
2E7
3
4
5
NCOA1
HDAC4
HIST2H2BE
TSPYL4
ATP10A
DLX5
DCN
SLC22A18
EPAS1
SNRPN
SNRPN
NCOA3
SMARCA2
HDAC8
HMG20B
SET7
JMJD2C
SAT2
KIAA0907
SUV420H1
HDAC5
MECP2
SMARCE1
H3F3A
MLL3
CHD3
MYST4
NDN
JMJD2B
ZBTB7A
SIRT2
H2AFJ
H2AFY2
ZNF264
ZNF264
CDKN1C
NAP1L3
-3.0 -2.4 -1.8 -1.2 -0.6 0 0.6 1.2 1.8 2.4 3.0

Figure 4.10

b.

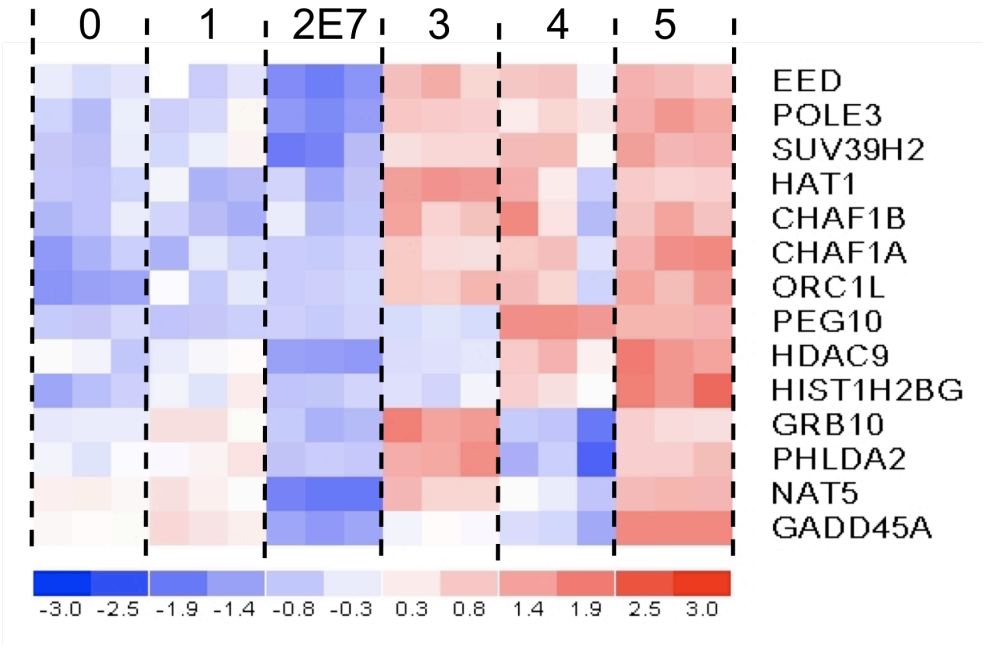


Figure 4.10: Expression of epigenetic regulator genes changes during step-wise transformation when corrected for cell proliferation. Gene expression microarray data for epigenetic regulators was normalized to expression of the proliferation-linked gene PCNA and analysed for significant changes in expression between MSC 0 and MSC 5. Heatmaps showing triplicate gene expression microarray data for significantly **(a)** down-regulated and **(b)** up-regulated genes in each of the five transforming steps (MSC 0 - 5).

4.4 Discussion

Results from the previous chapter indicated that Sat2 and LINE1 elements were hypomethylated in the MSC 5 cell line compared to previous stages in step-wise transformation. Experiments conducted as part of this chapter aimed to investigate the time course of hypomethylation following H-Ras expression to determine whether this loss of methylation was due to an active or passive mechanism. These experiments also aimed to determine whether transformation could take place in the absence of genome-wide hypomethylation. Additionally, I aimed to investigate changes to the expression of epigenetic regulators and metabolic enzymes during step-wise transformation in order to identify alterations that could lead to the observed hypomethylation.

My investigation of cellular transformation by the conditional ER^{TM}-H-Ras aimed to address the time course of demethylation following H-Ras^{V12} expression. However, these experiments did not demonstrate hypomethylation of Sat2 or LINE1 elements in transformed cells grown over the course of one month. Conversely, MSC 4 cells expressing H-Ras^{V12} without a conditional domain exhibited Sat2 hypomethylation when grown for the same length of time. Differences in the rate of methylation loss between the two systems may account for this observation. Alternatively, although the activity of the conditional ER^{TM}-H-Ras^{V12} was extensively characterised, differences between this fusion protein and the constitutively expressed H-Ras^{V12} could also affect the incidence of Sat2 hypomethylation. For example, although the fusion protein exhibited similar signaling properties through the MAPK cascade, downstream PI3K or Ral-GEF signaling may have been disrupted (Rangarajan et al., 2004).

Data obtained from both the conditional and the constitutively expressed H-Ras^{V12} cell lines point to a number of interesting conclusions. First, loss of Sat2 methylation following H-Ras^{V12} expression is a gradual event, reaching statistical significance only after four weeks. This suggests that repetitive element hypomethylation in this system does not result from an active demethylation process, which would be expected to induce hypomethylation in a shorter period of time. Although both active and passive demethylation have been shown to induce genome-wide hypomethylation in mouse cells, this is the first study to address the time course and potential mechanism of hypomethylation following transformation of human cells (Gaudet et al., 2003; Reik et al., 2001). Secondly, both the conditional and constitutively expressed H-Ras^{V12} cell lines exhibited transformation in the absence of repetitive element hypomethylation, suggesting that genome-wide hypomethylation is not a requirement of transformation.

Many other studies have attempted to address the question of whether genomic hypomethylation is causative of transformation or *vice versa*. For example, mice containing a hypomorphic allele of *dnmt1*, which reduces protein expression to 10% of wild type levels, exhibited global hypomethylation along with an increased incidence of T-cell lymphoma. The authors concluded that genome-wide loss of methylation resulted in transformation through increased incidence of karyotypic changes and chromosomal abnormalities (Gaudet et al., 2003). Although the causative role of hypomethylation during transformation could be tissue specific, studies in mice have suggested that soft tissue sarcoma incidence increases following global loss of methylation induced by *Dnmt1* hypomorphic alleles in *Nf1*$^{+/-}$ *p53*$^{+/-}$ mice (Eden et al., 2003). This finding is of

direct relevance to my study, since MSC are believed to be the cell of origin for various types of sarcoma (Helman and Meltzer, 2003). Whilst both of these reports suggest that genome-wide hypomethylation can promote transformation, they do not address whether transformation can take place in the absence of this loss of methylation.

However, another mouse model of genome-wide hypomethylation induced by *dnmt1* knockdown notes opposing effects on the growth of different tumour types (Yamada et al., 2005). Incidence of liver and intestinal carcinogenesis was monitored in $APC^{Min/+}$ mice in which *Dnmt1* function was reduced using hypomorphic alleles, leading to increased development of multifocal liver tumours but decreased incidence of macroscopic intestinal tumours. The authors speculate that transformation events requiring loss of imprinting benefited from genome-wide hypomethylation, whereas this loss of methylation was detrimental to transformation processes dependent on epigenetic gene silencing by promoter methylation (Yamada et al., 2005). Together, these reports show that genome-wide hypomethylation can promote some instances of carcinogenesis but there is no definitive demonstration that hypomethylation is a strict requirement of transformation.

Similarly, the timing of genome wide hypomethylation during the development of cancer could provide an indication of whether this change is required for transformation to take place. Many studies have suggested that genome-wide DNA hypomethylation is an early event in transformation, and can be observed in benign tumours or adjacent normal cells. Genome-wide hypomethylation is also included as an early event in a classic model of step-wise transformation in colorectal cancer (Fearon and Vogelstein, 1990). This

model cites two early investigations of genome-wide methylation levels in benign and malignant colonic neoplasms as evidence to this effect (Feinberg et al., 1988; Goelz et al., 1985). A more recent study of LINE1 promoter methylation (a marker of genome-wide methylation levels) in colorectal cancer patients observed genome-wide hypomethylation in tumours compared to tissue from healthy individuals. Furthermore, hypomethylation was also evident in adjacent normal tissues in 6 out of 19 cancer patients, but was not present in peripheral blood from the same individuals. The authors concluded that global hypomethylation in the adjacent normal tissue indicates that this change is an early event in the transformation process, potentially leading to karyotypic instability and further oncogenic changes (Suter et al., 2004).

However, several other investigations have failed to detect reduced genome-wide methylcytosine levels in benign cancers. One of the earliest studies of genome-wide methylation levels in cancer surveyed a range of benign and neoplastic tissues by HPLC (Gama-Sosa et al., 1983b). These data indicated that global methylation levels in benign tumours were very similar to those observed in normal tissues. Furthermore, a large proportion of malignant samples displayed genome-wide hypomethylation, with secondary malignancies exhibiting the greatest degree of hypomethylation (Gama-Sosa et al., 1983b). Several studies have also addressed global methylation levels in different stages of ovarian carcinogenesis. In one such investigation, satellite repeat methylation levels were assessed by genomic DNA digestion with methylation specific restriction enzymes followed by southern blot hybridization. This indicated that while ovarian carcinomas and low malignant potential tumours were significantly hypomethlyated compared to normal tissues, benign cystadenomas

displayed weak or no hypomethylation (Ehrlich et al., 2006). Together, these studies show that genome-wide hypomethylation is not always detectable in early stage or premalignant lesions, suggesting that this epigenetic change is not required for transformation to take place.

The relative timings of global hypomethylation and gene specific hypermethylation in cancer have also been utilized to address whether hypomethylation occurs in the early stages of transformation. Although many studies have failed to elucidate a link between the two events, an investigation of LINE1 methylation levels in prostate carcinomas of different stages found that samples exhibiting hypomethylation formed a subclass of those with gene specific hypermethylation. The lack of tumours exhibiting hypomethylation in the absence of gene specific hypermethylation suggests that global loss of methylation is a secondary event to hypermethylation in the progression of this cancer. This finding was supported by another study of DNA hypomethylation in prostate cancer progression, which noted significant hypomethylation only in metastatic cancers, later than CpG island hypermethylation (Yegnasubramanian et al., 2008; Yegnasubramanian et al., 2004).

Together, these studies suggest that although genome wide hypomethylation has traditionally been viewed as an early event in carcinogenesis, there is ample evidence supporting both this view and the alternative, that hypomethylation can occur later in transformation or not at all. Results obtained from my investigation suggest that hypomethylation can occur during step-wise transformation, although it is not a necessary event for transformation to take place. Furthermore, in the instance where

hypomethylation was observed, this was a late event in this step-wise genetic model, occurring on the introduction of the last oncogenic hit.

In order to explore the possibility that altered expression of epigenetic regulators caused repetitive element hypomethylation in MSC 5, an analysis of gene expression microarray data was conducted for each of the steps in the MSC model, with significant changes in the expression of key epigenetic regulators extracted from the data set. Furthermore, order to correct these observations for changes in gene expression resulting from differences in proliferation between cell lines, the analysis was repeated using expression values normalized against expression of the proliferation marker PCNA. The different results obtained from these two analyses suggest that many of the gene expression changes observed in the unadjusted array data may be caused by transcriptional changes linked to differences in the rate of proliferation between MSC cell lines, further highlighting the importance of this consideration when interpreting results obtained from the model. Therefore, the results obtained from the adjusted array data may give a stronger indication of true changes in expression caused by the genetic hits introduced in this model.

These data indicated that of the genes involved in the establishment and maintenance of DNA methylation, only *DNMT1* showed a significant change in expression, increasing between MSC 0 and 5 in the uncorrected array analysis. However, when adjusted for proliferation, expression of *DNMT1* does not change significantly during step-wise transformation. It remains possible, therefore, that the genome-wide hypomethylation observed during step-wise transformation of MSC could result from insufficient DNMT activity to cope with the faster cell cycle in the later stages of step-wise transformation.

However, future work on DNMT1 enzymatic activity should be conducted in order to establish whether any changes in enzymatic function contribute to hypomethylation in MSC 5.

Although many studies have assessed DNMT expression levels in cancer, no clear link between altered DNMT expression and the incidence of genome-wide hypomethylation has emerged. One study of genome-wide hypomethylation in uterine leiomyoma noted unchanged or elevated *DNMT1* expression and decreased *DNMT3A* and *DNMT3B* expression compared to adjacent normal tissue. Additionally, genome-wide hypomethylation in ovarian epithelial tumours of different malignant potential has been shown to be independent of RNA levels for all DNMTs, although no comparison of DNMT expression in normal tissues was conducted (Ehrlich et al., 2006). A similar investigation of Wilm's tumour showed global hypomethylation in 60% of samples when compared to a panel of normal postnatal tissues, but failed to find any correlation between global methylation levels and relative RNA levels of *DNMT1*, *DNMT3A* and *DNMT3B* (Ehrlich et al., 2002).

As well as passive demethylation processes, global hypomethylation could be induced in an active manner through increased expression or activity of DNA demethylases. Although active DNA demethylation has been observed during early mouse embryogenesis, a DNA demethylase in human cells has not yet been identified (Morgan et al., 2004; Ooi and Bestor, 2008). Analysis of gene expression microarray data for putative demethylases including *MBD2* and *GADD45A* revealed a significant increase in *GADD45A* expression in MSC 5 compared to MSC 0, with this increase occurring on the introduction of H-Ras^{V12}. Furthermore, this significant change in gene expression is also evident

when the gene expression microarray data is normalised for proliferation rate of the cells. It remains a possibility, therefore, that increased demethylase activity of this enzyme in transformed MSC could result in global hypomethylation, although future characterisation and studies of this enzyme in the MSC model are required to test this hypothesis.

Normal DNA methylation patterns may also be disrupted by changes in the availability of SAM, a key substrate required for the activity of DNMTs. Disrupted SAM biosynthesis through dietary alterations or disrupted methoionine metabolism has previously been linked to global hypomethylation and cancer (reviewed in Kim, 2005; Laird and Jaenisch, 1996; Szyf et al., 2004). Gene expression microarray data from the MSC model suggest that metabolic pathways directed away from SAM production are up-regulated in transformed MSC, raising the possibility that SAM levels are limiting in these cells. These observations are concurrent with other studies in which polymorphisms that reduce the activity of MTHFR, a key player in methyl-donor biosynthesis, are associated with increased cancer risk (reviewed in Szyf et al., 2004). Reduced SAM availability may therefore contribute to the hypomethylation observed in MSC 5.

Together, these data indicated that there is no single change in gene expression that could lead to repetitive element hypomethylation in MSC 5. However, concurrent changes to the expression of a number of different enzymes involved in DNA methylation suggested that the combinatorial effects of these alterations could result in a gradual loss of normal DNA methylation patterning over time.

In conclusion, experiments conducted as part of this chapter have aimed to address the time course and potential mechanism of repetitive element hypomethylation during step-wise transformation of MSC. Analysis of repetitive element methylation over time following conditional H-Ras^{V12} expression failed to detect hypomethylation of these sequences despite the evident transformation of these cells. Conversely, expression of H-Ras^{V12} lacking a conditional domain was able to recapitulate the Sat2 hypomethylation previously observed in the original MSC 5 population. This hypomethylation reached significance following growth of these cells for one month, suggesting that methylation is lost in a gradual manner, potentially through a passive mechanism. Finally, analysis of gene expression microarray data for epigenetic regulators and enzymes relating to methionine metabolism indicated deregulation of these pathways during transformation. Together, these results provide insight into the causes of genome-wide hypomethylation in cancer and suggest that this loss of methylation is not a requirement for transformation.

Chapter 5: PRC2 up-regulation and mechanisms of target gene silencing during step-wise transformation

5.1 Background

Although genome-wide hypomethylation is one of the most widely reported epigenetic changes in cancer, other hallmark epigenetic changes occur during transformation, including gene specific hypermethylation and characteristic histone modifications (Esteller, 2006; Fraga et al., 2005; Herman and Baylin, 2003; Jones and Baylin, 2002). An important class of epigenetic regulators involved in the latter two of these changes is the polycomb repressive complex 2 (PRC2) - an epigenetic silencing complex essential for normal developmental processes and commonly deregulated in cancer (Sparmann and van Lohuizen, 2006). The main action of PRC2 is to maintain embryonic stem cell pluripotency by repression of genes involved in differentiation (Lee et al., 2006). Gene silencing is achieved by creation of the repressive H3K27me3 histone modification, one half of the 'bivalent' chromatin state found at the promoter regions of key developmental genes in embryonic stem cells (Cao et al., 2002; Kuzmichev et al., 2002). This unique chromatin state maintains these genes in an unexpressed state, but allows them to be quickly activated upon differentiation (Bernstein et al., 2006).

Up-regulated expression of the PRC2 components EZH2 and SUZ12 has been well documented in a variety of cancers (Sparmann and van Lohuizen, 2006). In many cases EZH2 has been shown to act as an oncogene, with over-expression leading to increased cell proliferation, invasion of benign cells,

colony formation and *in vivo* tumour growth in nude mice (Bracken et al., 2003; Kleer et al., 2003; Varambally et al., 2002). Furthermore, *EZH2* knockdown results in growth arrest of cancer cells as well as decreasing tumour growth and metastasis *in vivo* (Croonquist and Van Ness, 2005; Varambally et al., 2002). A number of mechanisms for elevated EZH2 expression in cancer have been identified, including disrupted microRNA expression and the action of viral oncoproteins

The repressive action of PRC2 suggests a role for this complex in gene silencing during tumourigenesis (Schlesinger et al., 2007). The identification of a 'stem cell like' signature of gene silencing in cancer has given weight to this hypothesis, with a number of studies demonstrating preferential hypermethylation and silencing of PRC2 targets in transformed cells (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). This suggests that aberrant PRC2 activity in cancer may be a key process in maintaining transformed cells in a 'stem cell-like' state by suppressing differentiation.

Although many studies have focused on changes in PRC2 expression and activity in cancer, none have investigated how or when these changes occur during transformation. Furthermore, although *EZH2* has been shown to have oncogenic properties in cancer cells, the role of PRC2 during early stages of transformation is unknown. Step-wise transformation of MSC provides an excellent model in which to investigate changes in PRC2 expression and subsequent downstream gene silencing, allowing insight into the mechanism and time scale of these changes to be gained. I hypothesized that alterations in the expression and activity of the PRC2 complex take place during step-wise transformation, and that these changes result in hypermethylation and silencing

of PRC2 target genes in transformed MSC. I therefore aimed to investigate the role of PRC2 during step-wise transformation of MSC.

5.2 Aims

To investigate the role of the PRC2 complex during step-wise transformation of MSC. Particular questions to address include:

1. Are PRC2 components up-regulated during step-wise transformation?
2. Are PRC2 target genes down-regulated during transformation?
3. Is this down-regulation a result of increased H3K27me3 or DNA methylation at target gene promoters?
4. Can cells be transformed in the absence of functional PRC2?

5.3 Results

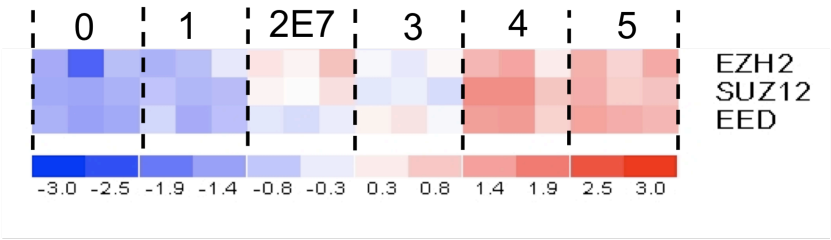
5.3.1 Analysis of Gene Expression Microarray Data and qRT-PCR Confirmation

In order to establish whether the three core components of PRC2 (EZH2, EED and SUZ12) are up-regulated during step-wise transformation, gene expression microarray data from the MSC cell lines were analysed. Changes in gene expression between primary and transformed MSC of greater than 1.5 fold with a p-value of less than 0.05 as determined by the Student's T-test were considered significant. This analysis revealed that *EZH2*, *EED* and *SUZ12* mRNA levels were significantly higher in MSC 5 compared to MSC 0 (Figure 5.1a; $q = 1x10^{-6}$, $q = 0$ and $q = 0$, respectively). To show that this up-regulation was specific to the PRC2 complex, gene expression data for PRC1 components were extracted and analysed (Figure 5.1b). These data confirmed that there was no significant change in expression for any of the genes involved with this alternate epigenetic regulator complex, suggesting that PRC2 is specifically up-regulated during step-wise transformation.

These findings were confirmed by qRT-PCR analysis of *EZH2*, *EED* and *SUZ12* mRNA levels in each of the MSC cell lines (Figure 5.1c-e). In all cases, significant up-regulation of gene expression was observed, with mRNA levels of *EZH2* increasing 20-fold, *SUZ12* 4-fold and *EED* 2.2-fold in MSC 5 compared to MSC 0. This increase relative to MSC 0 was first observed at MSC 1 for *EZH2*, MSC 3 for *SUZ12* and MSC 5 for *EED*. Together, these data confirmed that PRC2 components are specifically up-regulated during step-wise transformation.

Figure 5.1

a.



b.

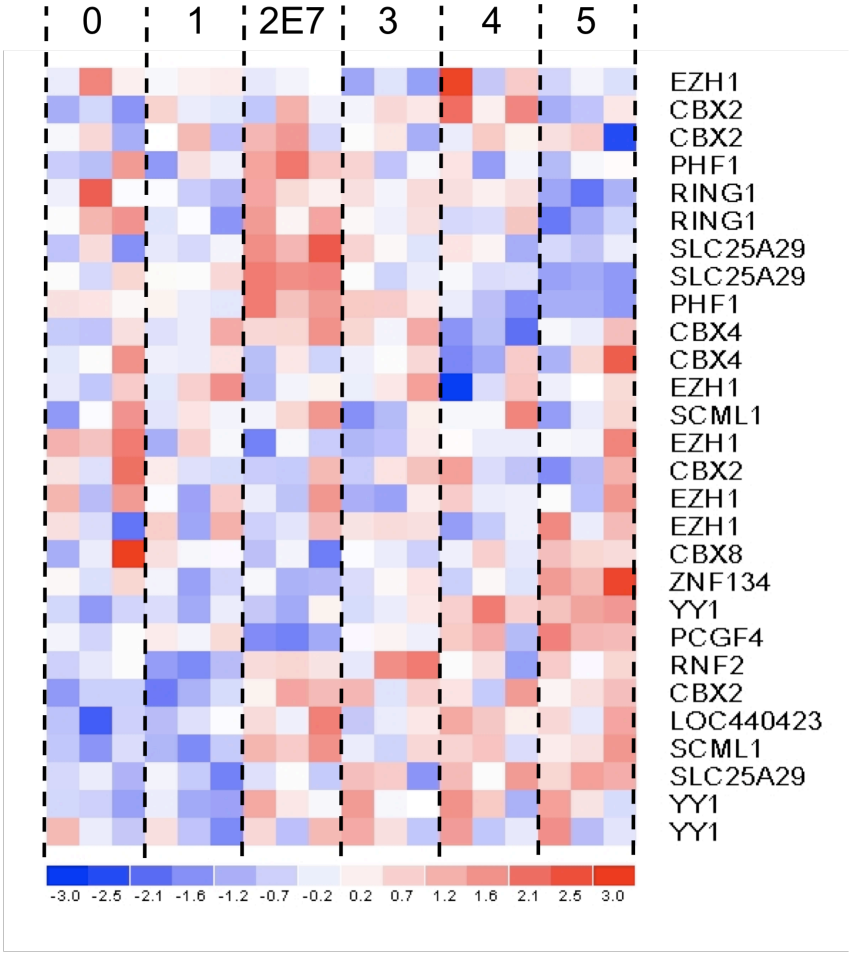


Figure continued on the following page

Figure 5.1

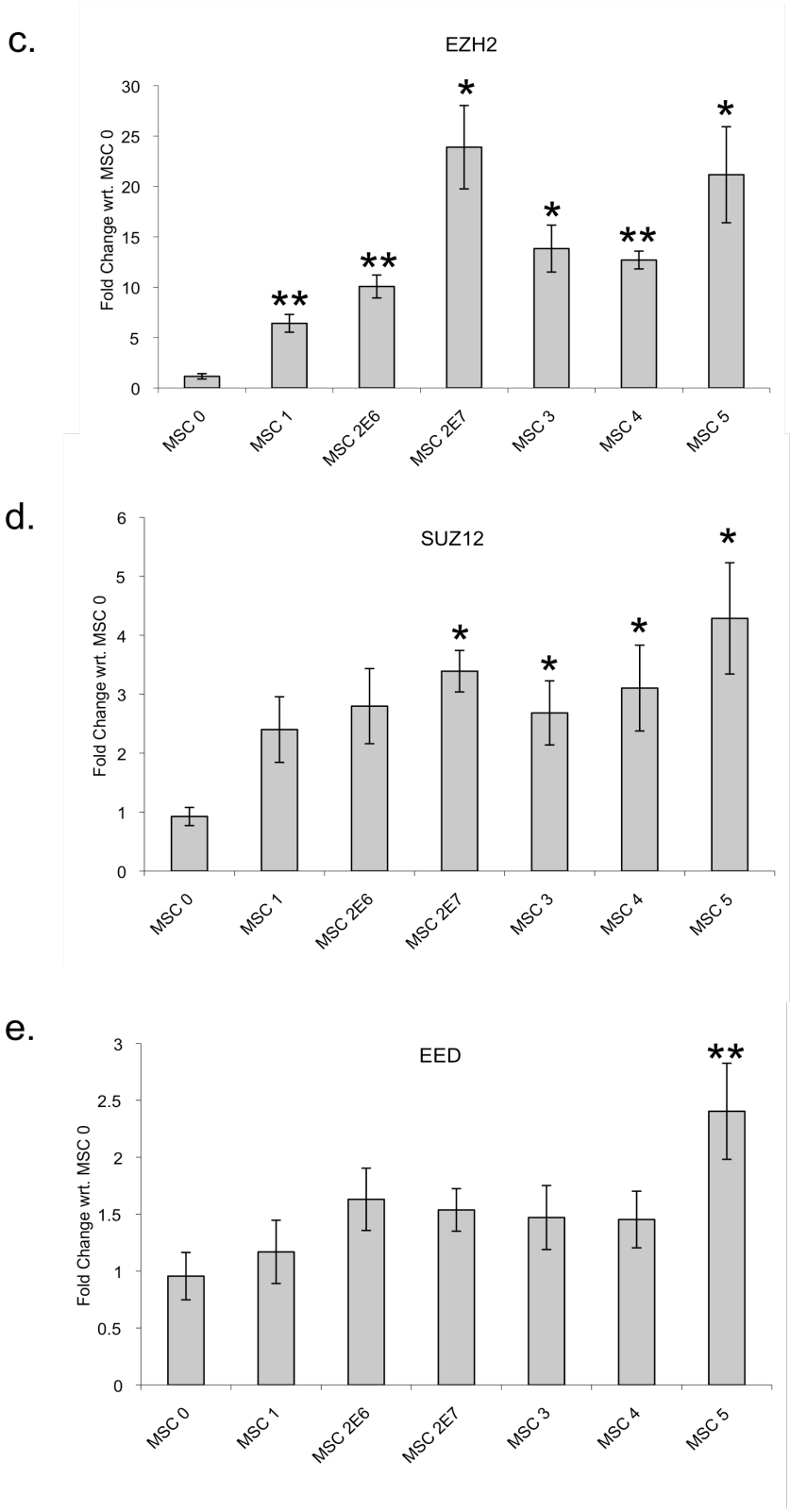


Figure 5.1: mRNA levels of the PRC2 components EZH2, EED and SUZ12 are up-regulated during step-wise transformation. Heatmap showing triplicate gene expression microarray data for **(a)** PRC2 and **(b)** PRC1 component expression in each of the five transforming steps (MSC 0 - 5). Expression profiles of **(c)** EZH2, **(d)** EED and **(e)** SUZ12 were confirmed by quantitative RT-PCR. Fold-change in mRNA levels with respect to MSC 0 were

calculated in triplicate by ΔΔCt using GAPDH as a loading control and normalised to parental MSC. Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$ with respect to MSC 0. Columns display average of three replicates; Error bars display SEM.

5.3.2 Western Blot Analysis

Western blot analysis was conducted to confirm that increased *EZH2*, *EED* and *SUZ12* mRNA levels translated to increased protein levels during step-wise transformation (Figure 5.2a). All MSC cell lines were grown in full media or serum starved for 24 hours to synchronize cells before total protein was harvested. Analysis of cell lysates revealed an increase in EZH2 and SUZ12 protein levels in both untreated and serum starved cells. No significant change in EED protein levels were observed in untreated cells although a modest increase was observed in serum starved cells from MSC 2E6 onwards.

To determine whether PRC2 up-regulation affected the amount of H3K27me3 in these cells, levels of the modified histone and total histone H3 were measured by western blot (Figure 5.2a). Although H3K27me3 levels fluctuated during step-wise transformation, changes to the amount of total H3 were also evident. To elucidate the underlying trend, densitometry of bands from triplicate western blot was conducted for each cell line, with the ratio of H3K27me3 : total H3 calculated (Figure 5.2b). For cells grown in full media, this analysis indicated that with the exception of MSC 1, there was no significant change in the ratio of H3K27me3 : total H3 in any of the MSC cell lines. In serum starved cells, however, this ratio was significantly lower in MSC 1 and at the later stages of transformation compared to MSC 0 (Figure 5.2b). These data suggested that genome-wide levels of H3K27me3 decreased during step-wise transformation of MSC, despite increased PRC2 expression.

Figure 5.2

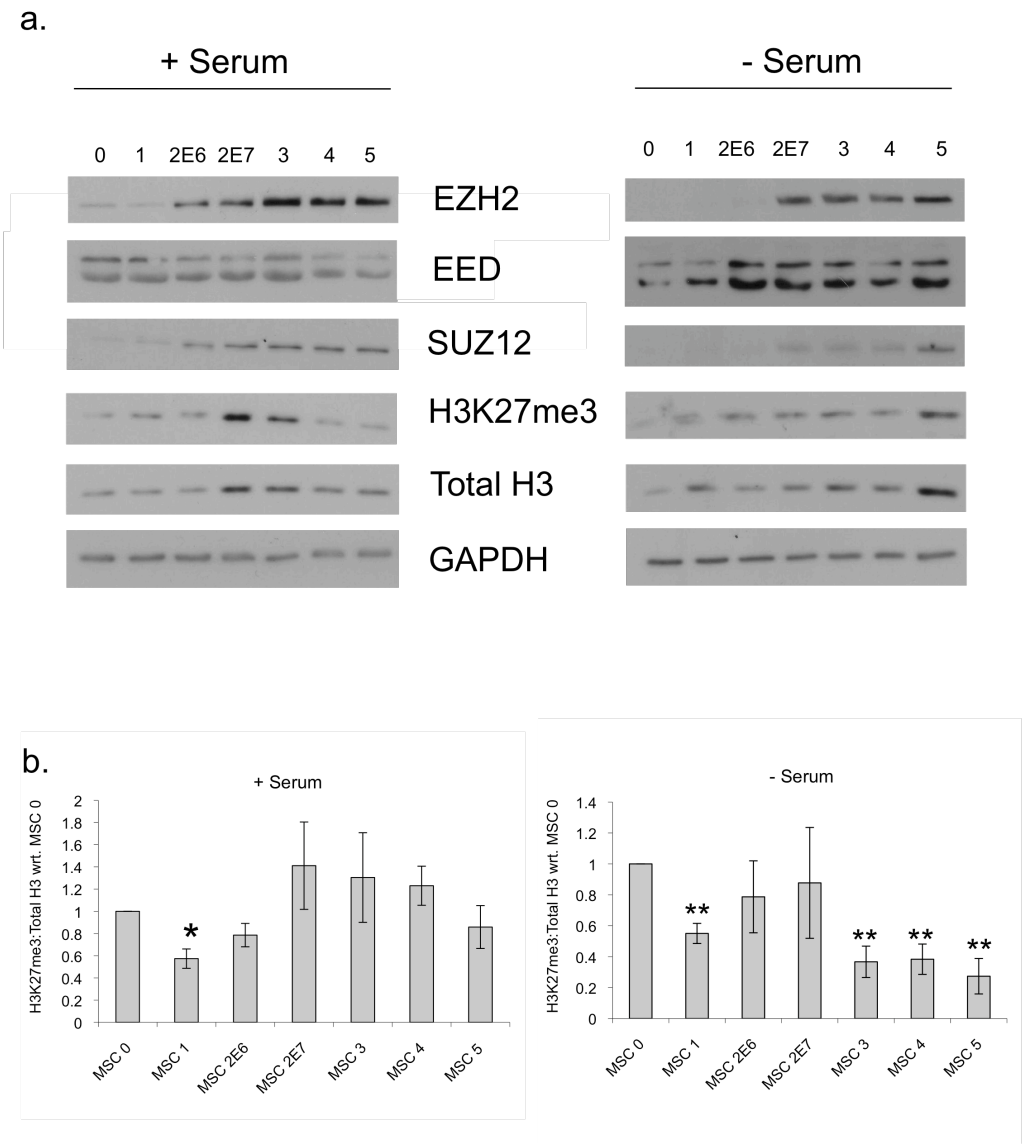


Figure 5.2: Genome-wide H3K27me3 levels are reduced during step-wise transformation despite up-regulation of PRC2. (a) Western blot analysis of EZH2, EED, SUZ12, H3K27me3, Total H3 and GAPDH (loading control) protein levels in MSC cell lines grown in full media (+ Serum) or serum starved for 24 hours (- Serum). **(b)** Quantification of the H3K27me3 : Total H3 ratio for all MSC cell lines grown in full media (+ Serum) or serum starved for 24 hours (- Serum). Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$ with respect to MSC 0. Columns display average of three replicates; Error bars display SEM.

5.3.3 *Gene Set Enrichment Analysis (GSEA) of PRC2 Target Gene Expression*

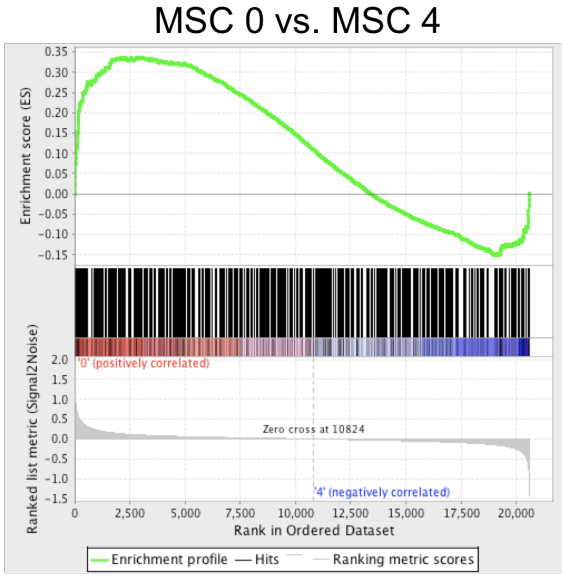
Since PRC2 acts as repressor complex, down-regulation of target gene expression would be expected following up-regulation of the complex in this model. Therefore, as another measure of PRC2 function, gene set enrichment analysis (GSEA) was conducted for PRC2 target genes. This unbiased computational approach compares two sets of gene expression microarray data, and determines whether a particular set of genes are significantly up- or down-regulated compared to the genome as a whole.

The set of PRC2 target genes utilized in my analysis were originally identified by chromatin immunoprecipitation followed by microarray hybridization (ChIP-chip) for SUZ12, EED and H3K27me3 in human embryonic stem cells (Lee et al., 2006). A set of 653 genes showing occupancy for all three of these marks were used for GSEA (Table A5). Comparison of MSC 0 to all other MSC cell lines indicated that PRC2 target genes were significantly enriched in MSC 0 compared to MSC 4 ($q = 0.001$), or in other words that these genes were significantly down-regulated at MSC 4 (Figure 5.3a). A significant enrichment score for this gene set was also obtained when MSC 3 was compared to MSC 4 ($q = 0$), confirming that significant changes to the expression of PRC2 target genes occurred at this stage of step-wise transformation. No significant enrichment was observed when MSC 0 was compared to any of the other cell lines, although the comparison of MSC 0 to MSC 5 approached statistical significance ($q = 0.069$). These results indicated that PRC2 target genes were significantly down-regulated during step-wise transformation, with this change in expression occurring at MSC 4.

GSEA also provides a list of 'core genes' that show the most significant changes in expression, and primarily account for the observed enrichment in the gene set. Core gene lists from GSEA of MSC 0 vs. MSC 4 and MSC 3 vs. MSC 4 were obtained, and expression data for genes contained in both lists were extracted (Tables 5.1 & 5.2). Analysis of these data confirmed the down-regulation of core gene expression between MSC 0 and MSC 4 (Figure 5.3b). A selection of candidate genes (DKK2, EN1, HOXB3, RBP4 and TBX2) with high GSEA core ranking and significant down-regulation in expression were selected for further investigation. Analysis of mRNA levels was conducted to confirm reduced expression of these genes between MSC 0 and MSC 4 (Figure 5.3c). These data displayed the expected pattern of down-regulation for all genes except HOXB3, which showed no significant change in expression between MSC 0 and MSC 4.

Figure 5.3

a.



b.

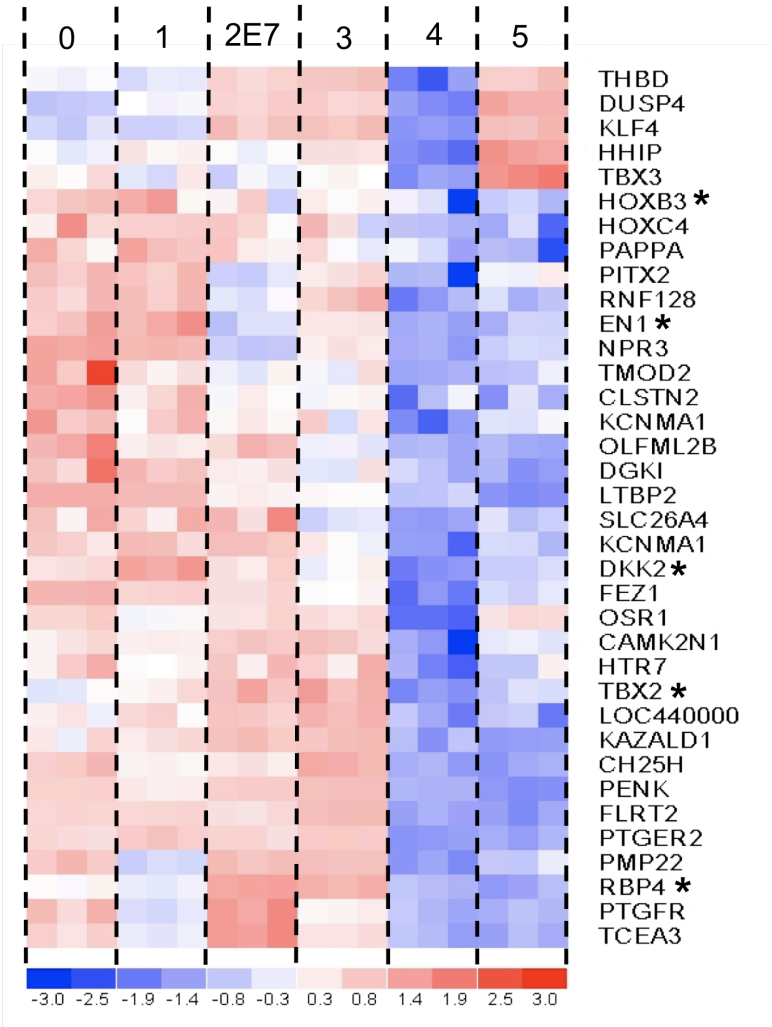


Figure continued on the following page

Figure 5.3

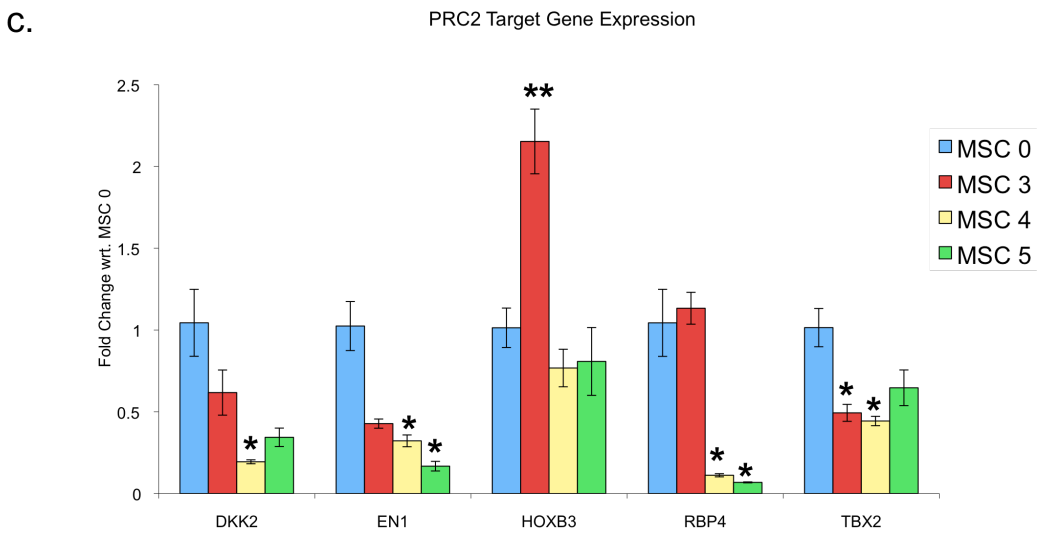


Figure 5.3: Gene set enrichment analysis (GSEA) indicates significant down-regulation of PRC2 target gene expression in MSC 4 versus MSC 0. (a) Enrichment plot from GSEA of PRC2 target gene expression profiles in MSC 0 vs. MSC 4, revealing significant down-regulation of target genes in MSC 4 compared to MSC 0. **(b)** Heatmap showing triplicate gene expression microarray data for GSEA core gene expression in each of the five transforming steps (MSC 0 - 5). Genes chosen for subsequent analysis are indicated by asterisks (*). **(c)** Quantitative RT-PCR confirmation of selected GSEA core gene down-regulation. Fold-change in mRNA levels of DKK2, EN1, RBP4, HOXB3 and TBX2 with respect to MSC 0 were calculated in triplicate by ΔΔCt using GAPDH as a loading control and normalised to parental MSC. Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$ with respect to MSC 0. Columns display average of three replicates; Error bars display SEM.

Table 5.1: Core gene list from GSEA of PRC2 target gene expression between MSC 0 and MSC 4

Gene Symbol	Rank In Gene List	Rank Metric Score	Running ES	Core Enrichment
OSR1	4	1.510	0.025	Yes
DKK2	22	1.092	0.042	Yes
FEZ1	28	1.075	0.059	Yes
RNF128	41	0.918	0.074	Yes
OLFML2B	56	0.841	0.087	Yes
PENK	65	0.795	0.100	Yes
CAMK2N1	72	0.777	0.112	Yes
SLC26A4	99	0.691	0.122	Yes
CH25H	124	0.643	0.132	Yes
PTGFR	126	0.642	0.142	Yes
TCEA3	137	0.623	0.152	Yes
NPR3	154	0.592	0.161	Yes
TMOD2	155	0.591	0.171	Yes
RBP4	163	0.575	0.180	Yes
FLRT2	171	0.564	0.189	Yes
BHLHB3	180	0.556	0.198	Yes
EN1	192	0.540	0.206	Yes
PTGER2	222	0.506	0.213	Yes
LTBP2	229	0.501	0.221	Yes
THBD	250	0.479	0.228	Yes
DGKI	327	0.417	0.231	Yes
IRX5	359	0.397	0.236	Yes
ADRA2A	363	0.396	0.242	Yes
PITX2	369	0.393	0.248	Yes
PAPPA	445	0.354	0.250	Yes
MAFB	481	0.340	0.254	Yes
TBX2	501	0.335	0.259	Yes
BTG2	513	0.330	0.264	Yes

Gene Symbol	Rank In Gene List	Rank Metric Score	Running ES	Core Enrichment
VDR	553	0.317	0.267	Yes
CLSTN2	585	0.309	0.271	Yes
KIAA1199	588	0.307	0.276	Yes
PTPRU	763	0.263	0.271	Yes
KCNK2	775	0.261	0.275	Yes
ZFYVE28	788	0.259	0.279	Yes
KLF4	853	0.246	0.279	Yes
PMP22	939	0.232	0.279	Yes
PDGFRA	947	0.229	0.283	Yes
PIR	956	0.229	0.286	Yes
TMEM27	958	0.228	0.290	Yes
KCNMA1	972	0.226	0.293	Yes
KAZALD1	1031	0.219	0.293	Yes
HTR7	1079	0.213	0.295	Yes
HOXC12	1116	0.208	0.296	Yes
FOXG1B	1149	0.203	0.298	Yes
NRG1	1201	0.198	0.299	Yes
PITX1	1209	0.197	0.302	Yes
HOXC4	1211	0.197	0.305	Yes
TBX3	1301	0.188	0.303	Yes
ZIC1	1356	0.183	0.304	Yes
HOXB2	1368	0.182	0.306	Yes
RIPK3	1378	0.181	0.309	Yes
NEF3	1400	0.178	0.311	Yes
ZNF503	1453	0.175	0.311	Yes
HLX1	1515	0.170	0.311	Yes
NKX2-2	1551	0.167	0.312	Yes
ZFHX1B	1563	0.166	0.314	Yes
MLLT3	1568	0.166	0.316	Yes
CRLF1	1570	0.166	0.319	Yes
TSLP	1591	0.165	0.321	Yes

Gene Symbol	Rank In Gene List	Rank Metric Score	Running ES	Core Enrichment
GHR	1611	0.163	0.322	Yes
PAX6	1643	0.161	0.324	Yes
NRG2	1651	0.160	0.326	Yes
PTHLH	1668	0.159	0.328	Yes
IRX3	1672	0.159	0.330	Yes
HBA2	1710	0.156	0.331	Yes
null	1779	0.151	0.330	Yes
ZNF436	1812	0.149	0.331	Yes
LOC153684	1843	0.147	0.332	Yes
FLJ37440	1859	0.147	0.333	Yes
HTR2C	1888	0.145	0.334	Yes
WNT16	1963	0.140	0.333	Yes
COL24A1	2021	0.136	0.333	Yes
LHX2	2045	0.135	0.334	Yes
NTN1	2082	0.133	0.334	Yes
CORO6	2094	0.133	0.336	Yes
DUSP4	2152	0.130	0.335	Yes
HOXB3	2212	0.128	0.334	Yes
MSC	2221	0.128	0.336	Yes
FBXL8	2264	0.126	0.336	Yes
GABRA2	2445	0.119	0.329	Yes
HES7	2485	0.118	0.329	Yes
ALOX15	2487	0.118	0.331	Yes
FRMD3	2513	0.117	0.331	Yes
HPSE2	2526	0.117	0.333	Yes
CA10	2533	0.116	0.334	Yes
UCP1	2542	0.116	0.336	Yes
RRP22	2691	0.112	0.330	Yes
RASSF5	2700	0.112	0.332	Yes
PODN	2744	0.110	0.331	Yes
CD8A	2780	0.109	0.331	Yes

Gene Symbol	Rank In Gene List	Rank Metric Score	Running ES	Core Enrichment
SIX1	2794	0.108	0.333	Yes
CDH7	2798	0.108	0.334	Yes
ASTN2	2810	0.108	0.335	Yes
GALNTL4	2811	0.108	0.337	Yes
HHIP	2852	0.106	0.337	Yes
FIGLA	2878	0.106	0.337	Yes

Table 5.2: Core gene list from GSEA of PRC2 target gene expression between MSC 3 and MSC 4

Gene Symbol	Rank In Gene List	Rank Metric Score	Running ES	Core Enrichment
OSR1	2	1.438	0.025	Yes
RBP4	20	1.069	0.042	Yes
RNF128	33	0.950	0.058	Yes
PTHLH	37	0.924	0.073	Yes
PENK	43	0.895	0.089	Yes
CAMK2N1	46	0.865	0.103	Yes
DKK2	48	0.854	0.118	Yes
THBD	53	0.827	0.132	Yes
BHLHB3	57	0.805	0.146	Yes
CRLF1	59	0.795	0.159	Yes
NEF3	66	0.734	0.171	Yes
CH25H	70	0.722	0.184	Yes
FEZ1	71	0.721	0.196	Yes
FLRT2	88	0.664	0.207	Yes
TBX2	91	0.658	0.218	Yes
KLF4	103	0.619	0.228	Yes
TCEA3	125	0.563	0.237	Yes
PTGER2	130	0.559	0.246	Yes
DUSP4	150	0.525	0.254	Yes
CYP24A1	162	0.497	0.262	Yes
EN1	242	0.399	0.265	Yes
NPR3	247	0.393	0.271	Yes
NPTX1	258	0.384	0.278	Yes
HHIP	265	0.379	0.284	Yes
PTGFR	266	0.377	0.290	Yes
KAZALD1	280	0.366	0.296	Yes
INA	284	0.362	0.302	Yes
FRMD3	300	0.353	0.307	Yes

Gene Symbol	Rank In Gene List	Rank Metric Score	Running ES	Core Enrichment
TSLP	312	0.348	0.313	Yes
PITX2	340	0.326	0.317	Yes
OLFML2B	344	0.324	0.322	Yes
SLC26A4	387	0.312	0.326	Yes
TMOD2	441	0.280	0.328	Yes
FLJ33790	467	0.268	0.331	Yes
WNT16	534	0.249	0.332	Yes
PITX1	543	0.247	0.336	Yes
PMP22	590	0.235	0.338	Yes
LTBP2	594	0.235	0.341	Yes
PAX6	647	0.222	0.343	Yes
HTR7	693	0.213	0.344	Yes
TRH	764	0.199	0.344	Yes
TLX2	864	0.182	0.342	Yes
PAPPA	924	0.174	0.342	Yes
DGKI	946	0.170	0.344	Yes
KCNMA1	947	0.170	0.347	Yes
NAGS	958	0.169	0.349	Yes
DLX2	986	0.165	0.351	Yes
HOXB3	1019	0.161	0.352	Yes
LHX2	1049	0.159	0.353	Yes
STMN2	1063	0.158	0.355	Yes
RGS20	1159	0.150	0.353	Yes
NPAS1	1211	0.146	0.353	Yes
TBX3	1218	0.145	0.355	Yes
CLSTN2	1234	0.143	0.357	Yes
ZIC1	1264	0.141	0.358	Yes
KCNC4	1271	0.140	0.360	Yes
EGR3	1296	0.139	0.361	Yes
HOXC4	1300	0.139	0.364	Yes
ZFHX1B	1301	0.139	0.366	Yes

Gene Symbol	Rank In Gene List	Rank Metric Score	Running ES	Core Enrichment
CXCL16	1383	0.133	0.364	Yes
COMP	1399	0.132	0.366	Yes
PKP1	1416	0.131	0.367	Yes
F2R	1435	0.129	0.368	Yes
ALOX15	1460	0.128	0.369	Yes

5.3.4 H3K27me3 Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation was conducted to determine whether the promoter regions of DKK2, EN1 and HOXB3 contained the H3K27me3 histone modification and whether levels of this modified histone change during transformation. Chromatin from MSC 0 and MSC 5 was harvested and subjected to immunoprecipitation with anti-H3K27me3 and anti-H3 antibodies. H3K27me3 levels at genomic regions immediately upstream of the transcription start sites for DKK2, EN1 and HOXB3 were assessed by qPCR. A genomic region upstream of DCN, a non-PRC2 target gene, was used as a negative control. These data showed that DNA from the anti-H3K27me3 immunoprecipitation amplified significantly more efficiently than DNA from the rabbit IgG control immunoprecipitation, indicating that the IP was successful (Figure 5.4 a & b). Furthermore, H3K27me3 levels were significantly higher at PRC2 target genes than at the negative control gene, DCN. Between cell lines, H3K27me3 levels at PRC2 target genes were significantly higher in parental MSC 0 than in transformed MSC 5 (Figure 5.4 a & b). These data indicated that PRC2 target genes were marked with the H3K27me3 histone modification in MSC 0, and that levels of this mark declined during step-wise transformation in a similar manner to genome-wide H3K27me3 levels (Figure 5.2b).

Figure 5.4

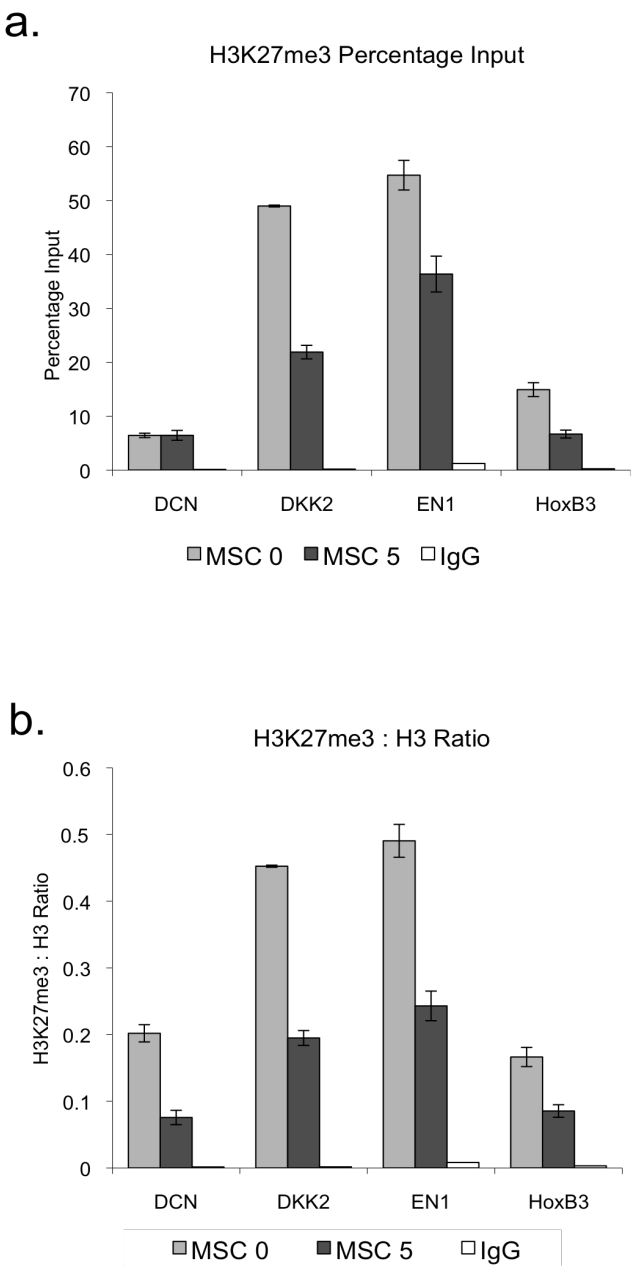


Figure 5.4: PRC2 target gene promoters are enriched for H3K27me3 in MSC 0, with levels of this modified histone decreasing during step-wise transformation. (a) Analysis of H3K27me3 and total histone H3 levels by ChIP at selected PRC2 target gene promoters and the non-PRC2 target gene DCN was conducted in MSC 0 and MSC 5. IP of MSC 5 chromatin with an equivalent amount of rabbit IgG was conducted as a negative control. H3K27me3 percentage of input for each cell line was calculated in triplicate relative to input reactions containing 100ng of template. **(b)** The ratio of percentage input for H3K27me3 and total H3 reactions were also calculated. Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$ with respect to MSC 0. Columns display average of three replicates; Error bars display SEM.

5.3.5 Analysis of Promoter DNA Methylation by Pyrosequencing

H3K27me3 has been hypothesized to direct DNA methylation and subsequent silencing of gene expression in cancer. To investigate whether DNA methylation at PRC2 target genes changes during step-wise transformation, pyrosequencing assays were designed for the promoter regions of DKK2, EN1, RBP4 and TBX2. CpG rich regions upstream of the transcription start sites of these genes were amplified by PCR from bisulphite converted genomic DNA and subjected to pyrosequencing. Methylation levels for each region were calculated as average percentage methylation across all CpG sites within the amplicon. These results indicated that the promoter regions of DKK2 and TBX2 were unmethylated in all MSC cell lines (Figure 5.5 a & d). Conversely, EN1 was moderately methylated in MSC 0, with significant hypermethylation occurring on further oncogenic hits up to MSC 4, before decreasing again in MSC 5 (Figure 5.5b). Finally, RBP4 showed low methylation in all MSC cell lines with the exception of MSC 4, where significant hypermethylation of the promoter was observed (Figure 5.5c).

Figure 5.5

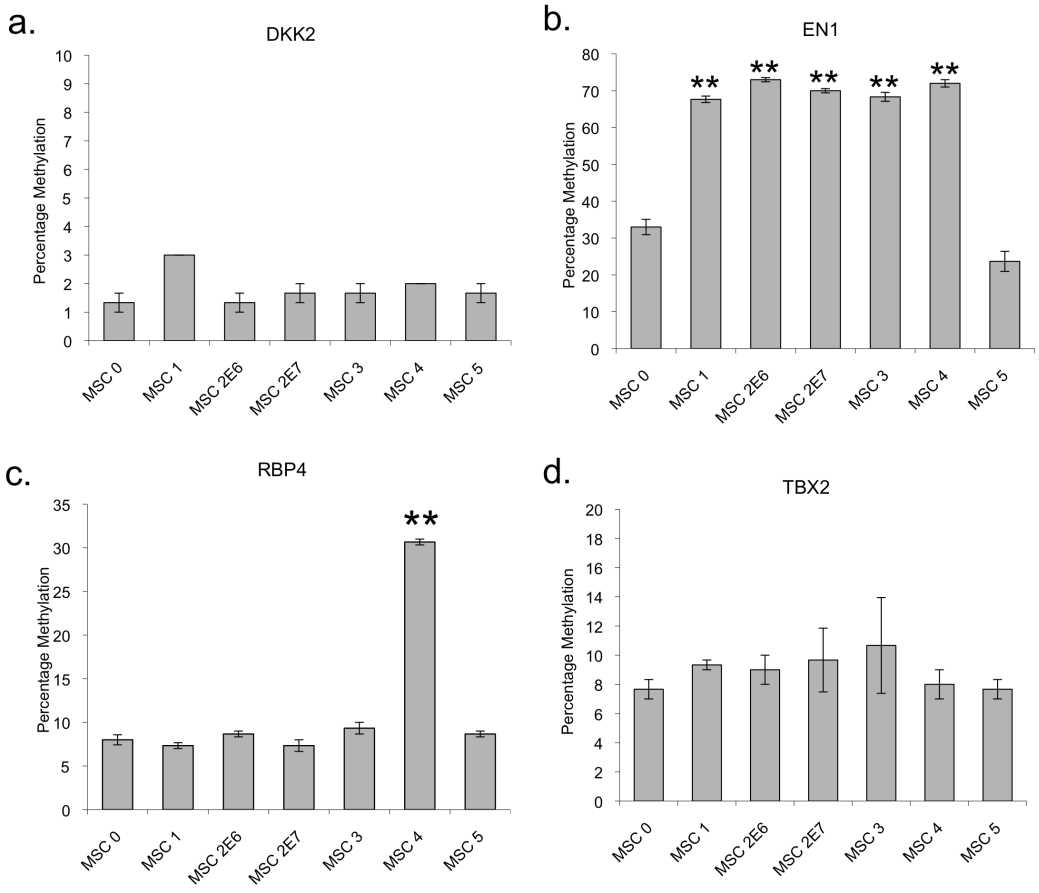


Figure 5.5: PRC2 target gene promoters show variable methylation patterns during step-wise transformation. Promoter regions from **(a)** DKK2, **(b)** EN1, **(c)** RBP4 and **(d)** TBX2 were amplified by PCR using bisulphite converted genomic DNA from all MSC cell lines (MSC 0 – 5) as well as from fully methylated and unmethylated control DNA (data not shown). CpG methylation in these amplification products was assessed by PyroSequencing in triplicate and averaged across the region assayed. Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$ with respect to MSC 0. Columns display average of three replicates; Error bars display SEM.

5.3.6 Short Hairpin RNA (shRNA) Knockdown of PRC2 Components

To investigate whether PRC2 target gene silencing in MSC 4 could be reversed, levels of EZH2 and SUZ12 were knocked-down in these cells using OpenBiosystems pGIPZ shRNA constructs (http://www.openbiosystems.com/). Stable gene knockdown was achieved by infecting target cells with lentivirus containing the shRNA expression construct. Since significant down-regulation of PRC2 target genes was observed in MSC 4, PRC2 knockdown was carried out in these cells, with the effect of this knockdown on PRC2 target gene expression assessed by qRT-PCR.

Lentiviruses containing shRNA constructs against *EZH2* and *SUZ12* along with non-silencing shRNA and empty vector controls were generated by FuGene transfection of 293T cells with the relevant lentiviral vectors. After 48 hours, supernatant containing the lentivirus was harvested and filtered. To control for the efficiency of viral production with different constructs, the lentiviral supernatants were titrated. Increasing volumes of supernatant were added to MSC 4, with the percentage of GFP cells assessed by FACS after 48 hours (Figure 5.6 a & b). In all cases, an increase in the percentage of GFP positive cells was observed on the addition of increasing volumes of lentivirus.

MSC 4 cells were infected with lentivirus containing either EZH2 shRNA, SUZ12 shRNA, non-silencing shRNA or empty vector. On the basis of viral titration, an appropriate volume of lentiviral supernatant was added to the cells in order to achieve a level of infection resulting in 30% GFP positive cells (Figure 5.6c). 48 hours post infection, GFP positive cells were sorted by MoFlow and cultured as normal (Figure 5.6d). Successful knockdown of EZH2 and SUZ12 relative to the non-silencing control was confirmed at the mRNA and

protein levels by qRT-PCR and western blot respectively (Figure 5.7a-c). Viability and growth of all cell lines was assessed using the MTS assay, indicating no significant decrease in growth potential for any of the shRNA expressing cells compared to uninfected MSC 4, although *EZH2* knockdown cells proliferated significantly faster than non-silencing control cells (Figure 5.7d; $p < 0.05$).

Figure 5.6

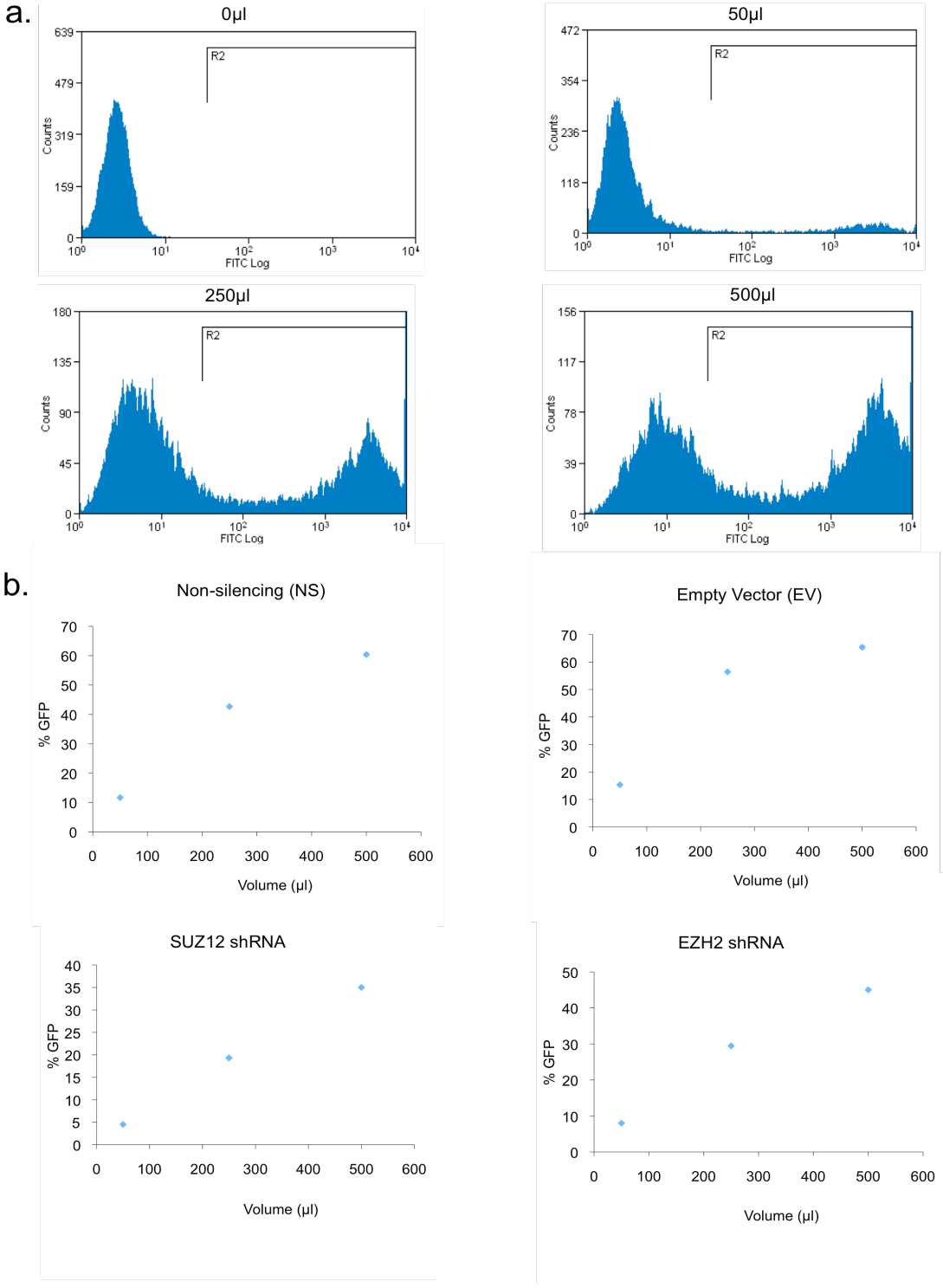

a.
0µl
50µl
250µl
500µl
R2
Counts
FITC Log
b.
Non-silencing (NS)
Empty Vector (EV)
SUZ12 shRNA
EZH2 shRNA
% GFP
Volume (µl)

Figure 5.6

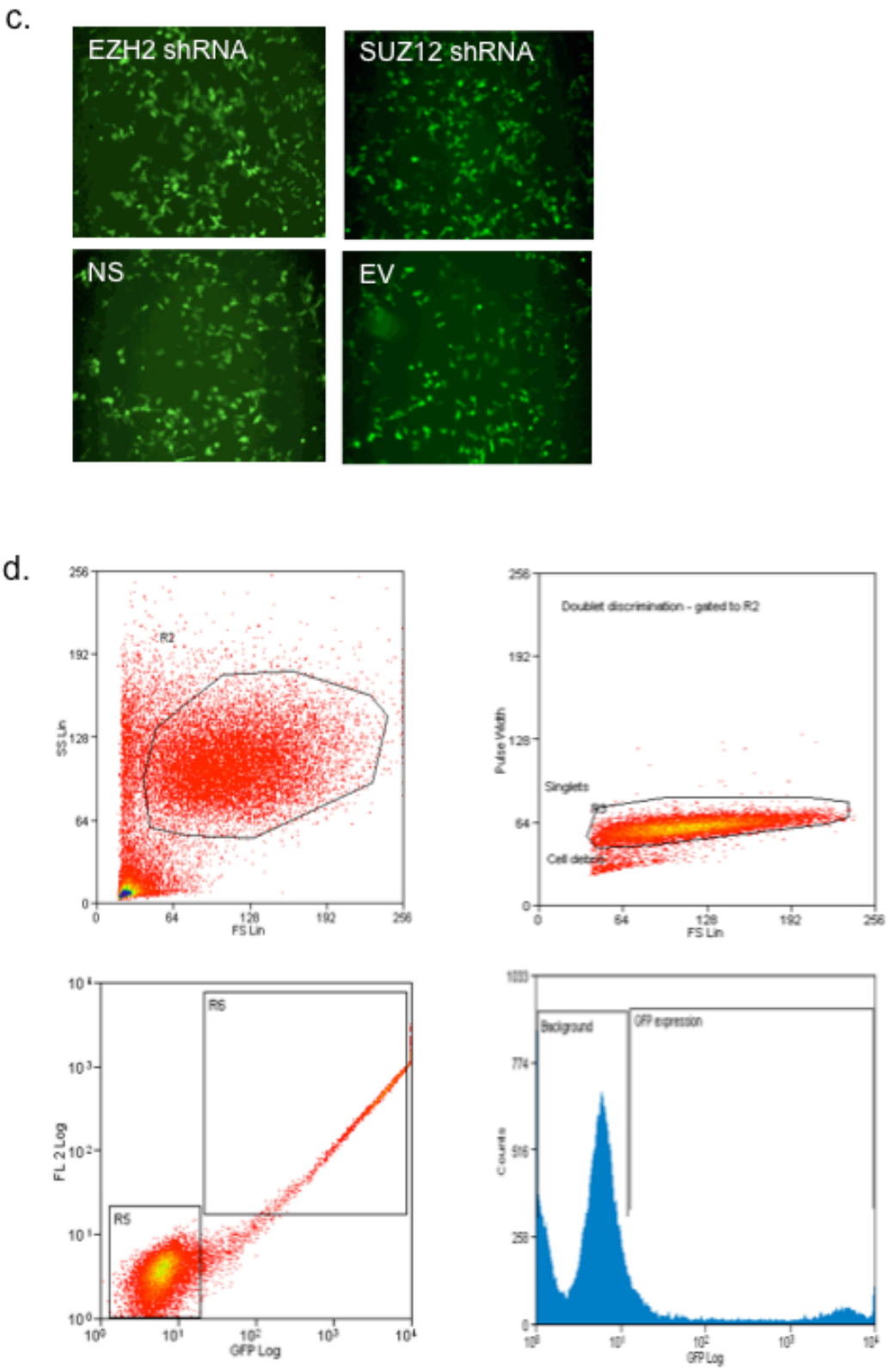


Figure 5.6: shRNA knockdown of PRC2 components in MSC 4. Lentiviruses containing shRNA constructs against *EZH2* and *SUZ12* along with non-silencing shRNA (NS) and empty vector (EV) controls were generated and used to infect MSC 4. **(a)** FACS plot displaying the GFP profile of MSC 4 cells following infection with increasing volumes of *EZH2* shRNA

lentivirus. **(b)** Analysis of FACS data for all lentiviral infections, displaying percentage of GFP positive cells following infection with increasing volumes of viral supernatant. **(c)** GFP image of MSC 4 cells infected with an appropriate volume of virus to achieve 10-25% GFP positive cells. **(d)** FACS plots from cell sorting of GFP positive cells infected with empty vector (EV) lentivirus.

Figure 5.7

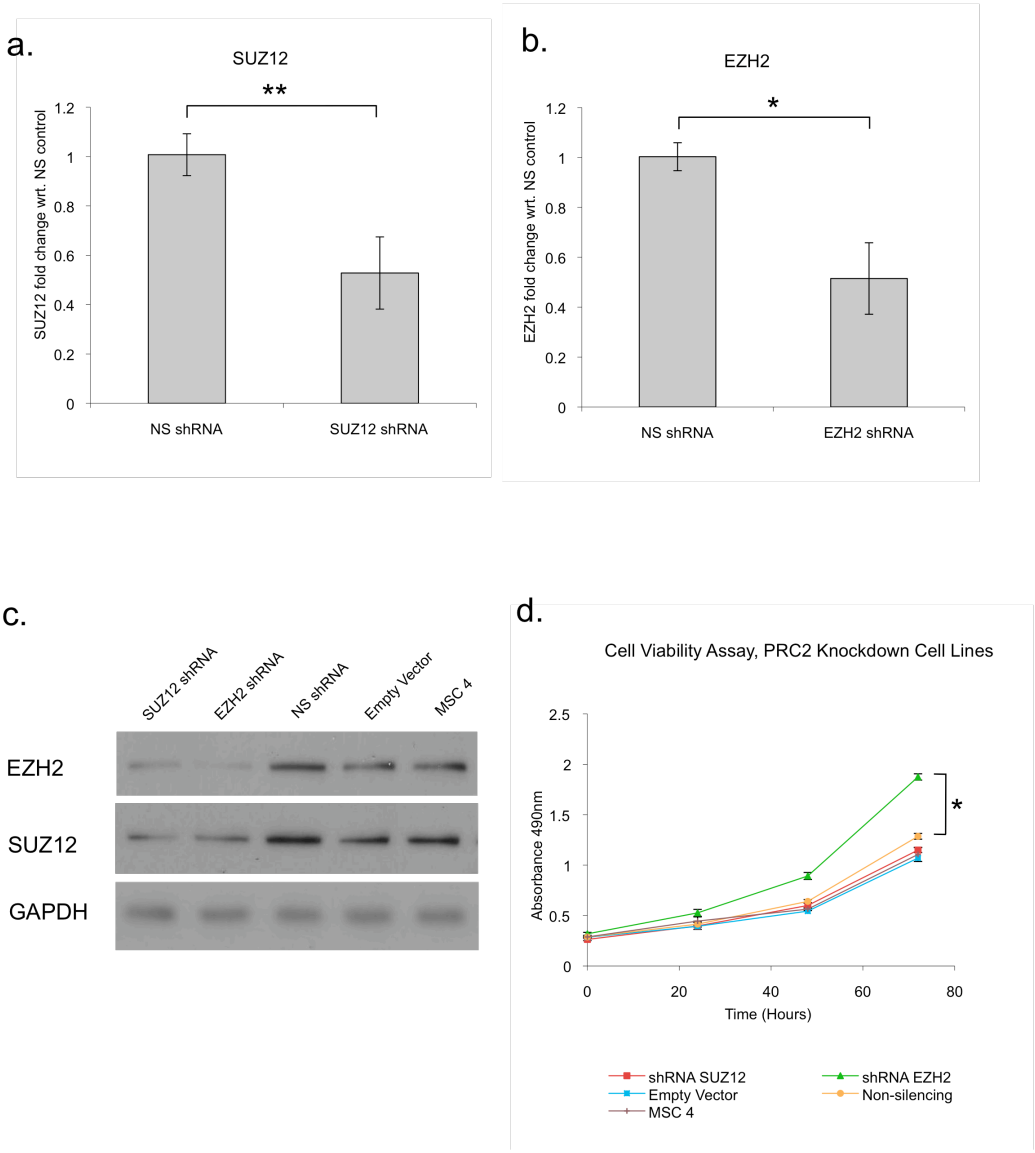


Figure 5.7: shRNAs induce knock-down of EZH2 and SUZ12 mRNA and protein in MSC 4. Quantitative RT-PCR confirmation of **(a)** EZH2 and **(b)** SUZ12 mRNA down-regulation in shRNA expressing cell lines. Fold-change in mRNA levels with respect to non-silencing shRNA control (NS) were calculated in triplicate by ΔΔCt using GAPDH as a loading control and normalised to NS. **(c)** Western blot analysis of EZH2, SUZ12 and GAPDH (loading control) protein levels in EZH2 and SUZ12 shRNA expressing cells, non-silencing shRNA expressing cells (NS), empty vector control cells (EV) and uninfected MSC 4 cells (MSC 4). **(d)** Cell viability (MTS) assay for all shRNA and control cell lines seeded at a density of 1000 cells per well and grown over the course of 72 hours. Each point represents the average absorbance (490nm) from 5 replicates when background absorbance is subtracted. Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$ with respect to MSC 0. Columns display average of three replicates; Error bars display SEM.

5.3.7 *PRC2 Target Gene Expression and Methylation following EZH2 and SUZ12 Knockdown*

To investigate whether down-regulation of PRC2 target gene expression was maintained in the absence of EZH2 and SUZ12, qRT-PCR for *DKK2*, *EN1*, *HOXB3*, *RBP4* and *TBX2* was conducted in the shRNA knockdown cell lines (Figure 5.8a). These data indicated no significant changes in expression for *EN1*, *HOXB3* or *RBP4* in the *SUZ12* or *EZH2* knockdown cell lines. Similarly, there was no change in *DKK2* or *TBX2* expression following *SUZ12* knockdown. However, a significant increase in expression of these genes was observed in *EZH2* knockdown cells compared to cells expressing a non-silencing control shRNA (Figure 5.8a).

Analysis of methylation at selected PRC2 target gene promoters was carried out by pyrosequencing in order to establish whether methylation levels changed following *EZH2* or *SUZ12* knockdown. Analysis of genomic regions upstream of the *DKK2*, *EN1*, *RBP4* and *TBX2* transcription start sites indicated no significant change in methylation following PRC2 knockdown (Figure 5.8b-e). Together, these data showed that expression of methylated genes (*EN1* and *RBP4*) was not affected by reduced *EZH2* expression, whereas unmethylated genes (*DKK2* and *TBX2*) were up-regulated in response to *EZH2* knockdown.

Figure 5.8

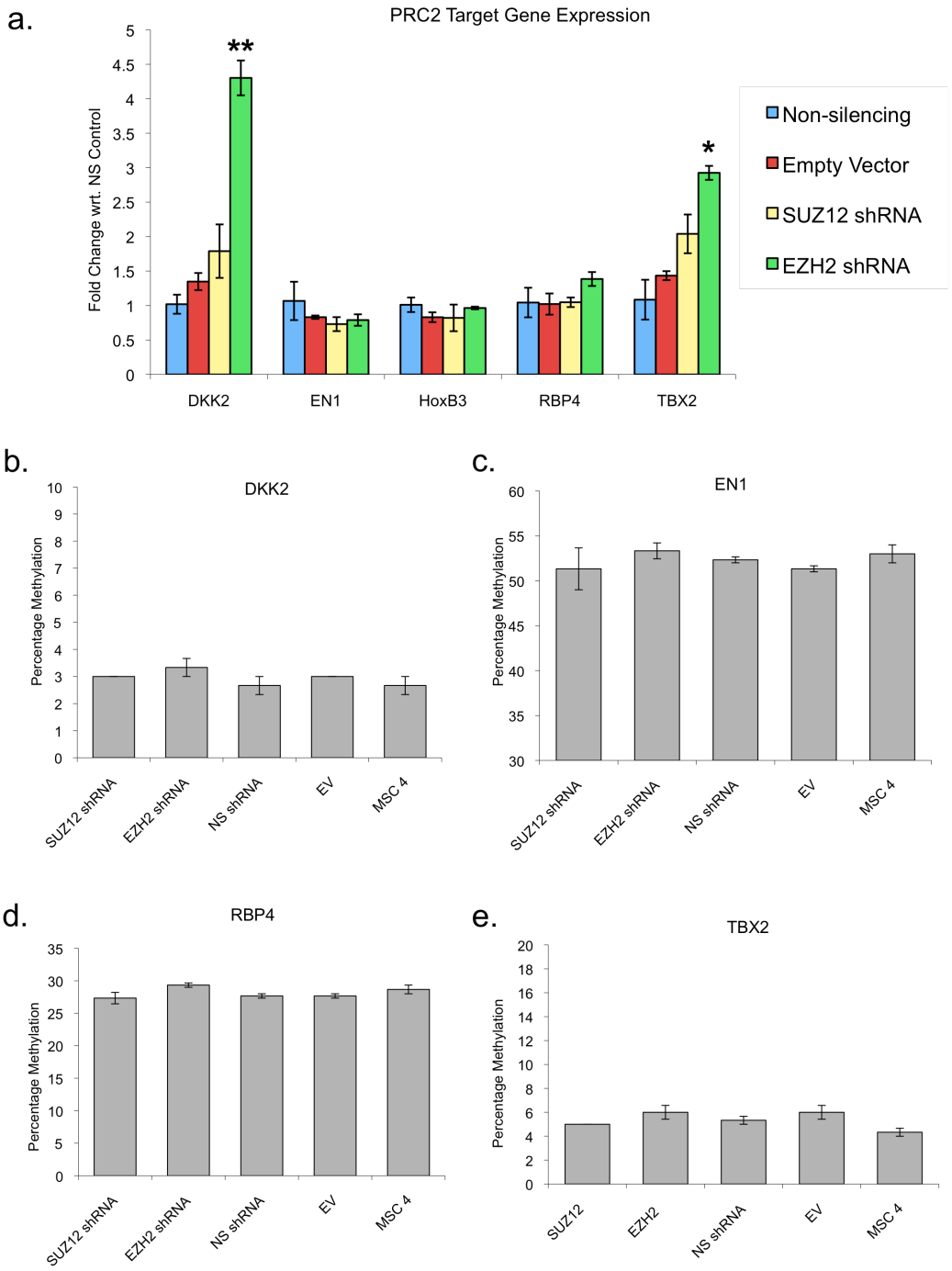


Figure 5.8: PRC2 target genes show differential changes in gene expression but no alteration in promoter DNA methylation following shRNA knockdown of *EZH2* or *SUZ12*. **(a)** Quantitative RT-PCR analysis of PRC2 target gene expression in shRNA expressing and control cell lines. Fold-change in *DKK2*, *EN1*, *RBP4*, *HOXB3* and *TBX2* mRNA levels with respect to non-silencing shRNA control (NS) were calculated in triplicate by ΔΔCt using GAPDH as a loading control and normalised to NS. **(b-e)** Average CpG methylation for **(b)** *DKK2*, **(c)** *EN1*, **(d)** *RBP4* and **(e)** *TBX2* promoter regions were assessed by Pyrosequencing in triplicate. Student's T-test was performed for all statistical analyses. * $p < 0.05$, ** $p < 0.01$ with respect to MSC 0. Columns display average of three replicates; Error bars display SEM.

5.3.8 H-Ras^{V12} Infection of PRC2 Knockdown Cells and Transformation Assay

In order to assess whether transformation could take place in the absence of PRC2 over-expression, H-Ras^{V12} was expressed in the *EZH2* and *SUZ12* knockdown cell lines. Retroviruses containing the H-Ras^{V12} construct or empty vector were generated and used to infect the MSC 4 knockdown cell lines. Following hygromycin selection of infected cells, H-Ras^{V12} expression and continued knockdown of EZH2 and SUZ12 was confirmed by western blot (Figure 5.9a). A viability assay of these cell lines indicated no significant difference in cell growth or viability between PRC2 knockdown and untreated MSC 4 cells expressing H-Ras^{V12} (Figure 5.9b). Finally, transformation of these cell lines was assessed by anchorage independent growth, with all cell lines able to form colonies when grown for 12 days in soft agarose (Figure 5.9c). Furthermore, colonies formed by the MSC 4 knockdown cells were GFP positive, indicating continued expression of the shRNA constructs. These data suggested that step-wise transformation could still occur in the absence of PRC2 over-expression.

Figure 5.9

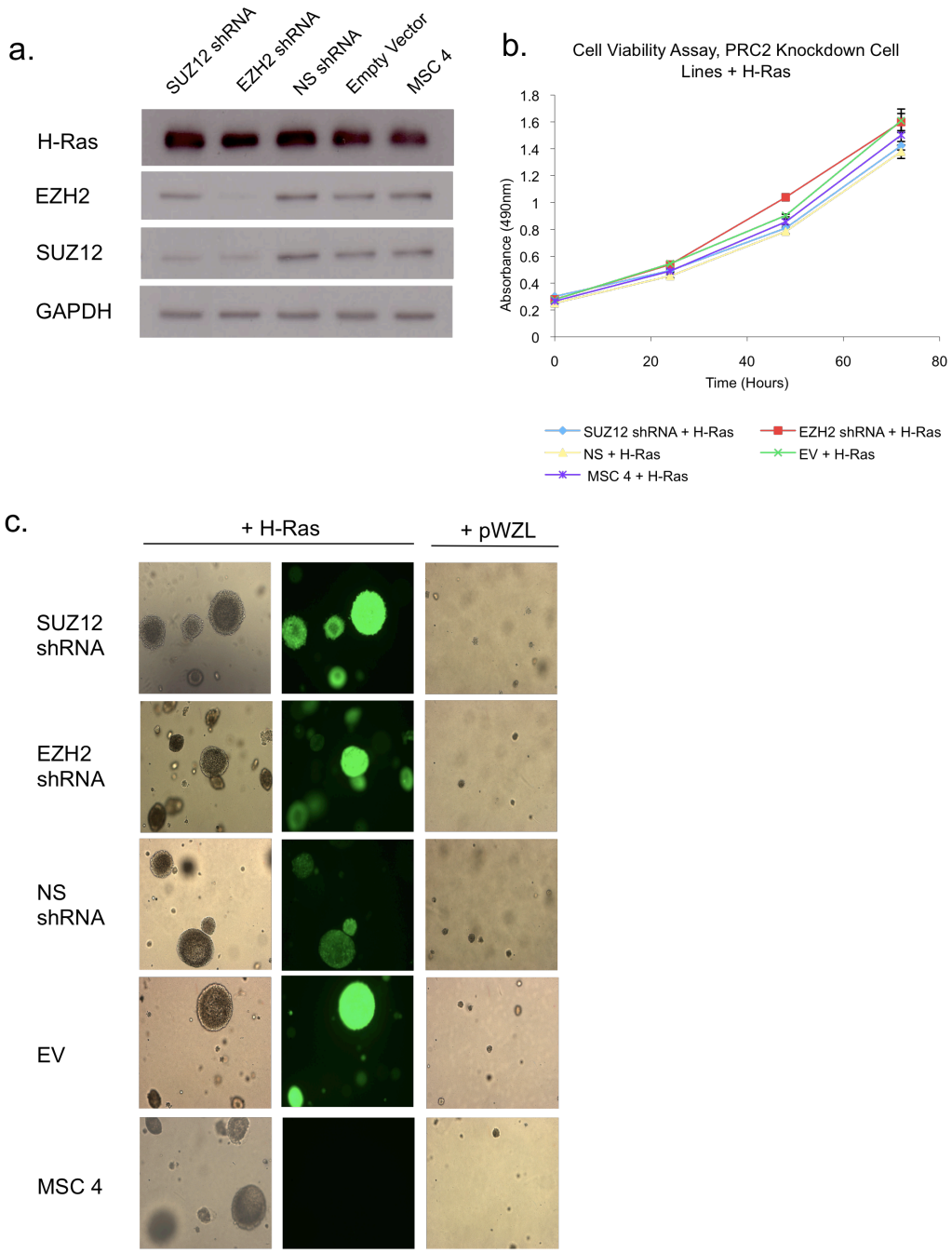


Figure continued on the following page

Figure 5.9

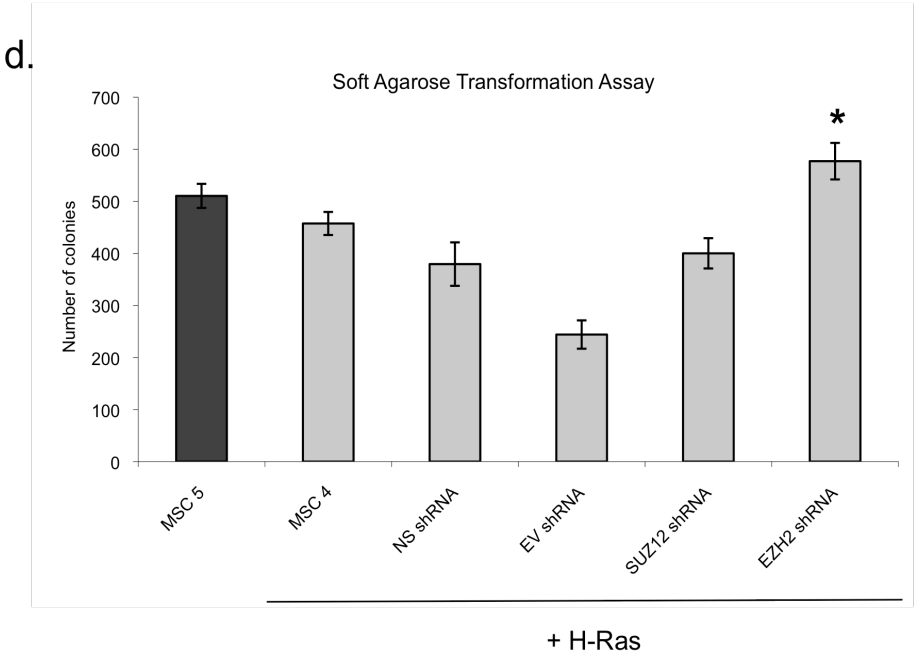


Figure 5.9: *EZH2* and *SUZ12* knockdown in MSC 4 does not prevent transformation on expression of H-Ras^{V12}. shRNA and control MSC 4 cell lines were infected with retrovirus containing a H-Ras^{V12} construct and subjected to selection with hygromycin. **(a)** Western blot analysis of H-Ras and GAPDH (loading control) protein levels following infection of MSC 4 cell lines with H-Ras^{V12} retrovirus. **(b)** Cell viability (MTS) assay for all shRNA and control cell lines following H-Ras^{V12} expression, seeded at a density of 1000 cells per well and grown over the course of 72 hours. Each point represents the average absorbance (490nm) from 5 replicates when background absorbance is subtracted. **(c)** Soft agarose transformation assay for all shRNA and control cell lines following H-Ras^{V12} expression or infection with empty vector control retrovirus (pWZL). Cells were seeded at a density of 10,000 cells per well and grown for 12 days. Positive control (MSC 5) and negative control (MSC 4) cells were grown in parallel (data not shown). Phase contrast and GFP images are depicted at 40x magnification. **(d)** Quantification of soft agarose colony numbers using SynGene GeneTools. Student's T-test was performed for all statistical analyses. * $p < 0.05$, wrt. NS shRNA + H-Ras. Columns display average of three replicates; Error bars display SEM.

5.4 Discussion

The PRC2 complex is an epigenetic silencing complex essential for the maintenance of embryonic stem cell pluripotency through repression of genes involved in differentiation (Boyer et al., 2006). Furthermore, this complex has been shown to play an important role in the development of cancer (Sparmann and van Lohuizen, 2006). PRC2 components are up-regulated in a variety of cancers, along with a characteristic epigenetic 'signature' of target gene silencing (Widschwendter et al., 2007). Although the importance of PRC2 over-expression in cancer has been well established, no investigation has addressed the role of this complex during transformation. The aims of this chapter were to further characterize changes in the expression and activity of PRC2 components during step-wise transformation and to establish whether transformation of MSC could take place in the absence of PRC2 over-expression.

Initial investigation of *EZH2*, *EED* and *SUZ12* mRNA and protein levels indicated up-regulation of all three components during step-wise transformation. This finding is consistent with previous studies of PRC2 expression, which have shown increased expression of *EZH2* and *SUZ12* in a variety of cancers (Sparmann and van Lohuizen, 2006). Furthermore, my data are consistent with previous models of *EZH2* regulation. For example, *EZH2* and *SUZ12* have previously been shown to be targets of the E2F transcription factor (Bracken et al., 2003; Weinmann et al., 2001). In line with this, there are significant increases in both *EZH2* and *SUZ12* mRNA levels when pRB activity is inhibited at MSC 2E7, leading to the release of E2F. However, although c-MYC has been shown to increase *EZH2* expression by way of miR-26a , no significant increase in *EZH2* mRNA is evident following the stabilization of c-MYC in MSC 4 .

A number of analyses were conducted in order to establish whether the increased expression of PRC2 components lead to functional changes in this model. Since PRC2 primarily acts to create the H3K27me3 histone modification, levels of this mark would be expected to increase in line with elevated PRC2 expression during transformation. Western blot analysis indicated changes in genome-wide levels of H3K27me3 during step-wise transformation. However, when these changes were corrected for alterations in the amount of total histone H3, a downward trend was evident, suggesting that H3K27me3 levels declined in the latter stages of step-wise transformation. This trend was also observed in the gene specific analysis of H3K27me3 levels by ChIP, which indicated lower levels of the modified histone at PRC2 target gene promoters in MSC 5 compared to MSC 0. Although no other studies have investigated global H3K27me3 levels during carcinogenesis in this way, altered EZH2 substrate specificity has been reported in cancer (Kuzmichev et al., 2004). EZH2 can form alternate PRC3 and PRC4 complexes with different EED isoforms, which act to methylate H1K26 rather than H3K27 (Kuzmichev et al., 2004). It is possible, therefore, that altered EZH2 substrate specificity in MSC 5 results in lower genome-wide levels of H3K27me3.

Altered PRC2 target gene preference in different cell types has also been suggested by a study of SUZ12 occupancy in a variety of normal and transformed human and mouse cell lines by ChIP-chip (Squazzo et al., 2006). Results from this investigation indicated that SUZ12 binding preference differed between tissues derived from embryonic or adult cells. However, a comparison of normal and transformed human cells was not conducted in this study, making it difficult to determine whether a similar shift in target gene preference occurs

during transformation. It is possible, therefore, that altered PRC2 target preference may contribute to the decline in H3K27me3 levels at PRC2 target genes between MSC 0 and MSC 5.

As another measure of PRC2 function, target gene expression profiles were analysed to determine whether repression of these genes occurred during step-wise transformation of MSC. GSEA indicated significant down-regulation of PRC2 targets in MSC 4 compared to MSC 0. Reduced expression of PRC2 target genes in cancer has also been suggested by ChIP-chip experiments in mouse F9 teratocarcinoma cells, indicating an inverse correlation between SUZ12 and RNA polymerase II occupancy at gene promoters (Squazzo et al., 2006).

To determine the mechanistic basis of this gene silencing, H3K27me3 and DNA methylation levels at PRC2 target gene promoters were analysed by ChIP and pyrosequencing respectively. These results showed that two PRC2 target genes – DKK2 and EN1 – were enriched for H3K27me3 in MSC 0 and MSC 5 compared to the non-PRC2 target gene DCN. However, levels of this modified histone at the promoter regions of these genes decreased between MSC 0 and MSC 5. Furthermore, HOXB3 – a well described target of the PRC2 complex in embryonic stem cells - did not show significant H3K27me3 enrichment. Expression of this gene also showed no decline during step-wise transformation, suggesting that it may not be targeted by the PRC2 complex in this cell type.

Methylation patterns at PRC2 target gene promoters also varied significantly. *DKK2* and *TBX2* showed minimal promoter methylation at all stages of step-wise transformation, whereas *EN1* and *RBP4* displayed

hypermethylation in some MSC cell lines, most notably MSC 4. However, the four promoter regions assayed in this study did not display hypermethylation in MSC 5 compared to MSC 0, contrary to previous reports of PRC2 target gene hypermethylation in cancer (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). These data indicated that hypermethylation of PRC2 target gene promoters is not always necessary for down-regulation of gene expression during transformation of MSC. However, it is possible that these genes remain silenced in MSC 5 through mechanisms other than DNA hypermethylation, such as repressive chromatin states. Furthermore, the genetic hits utilized in this model may have introduced a degree of redundancy, leading to gene silencing through other mechanisms and replacing the requirement for continued repression by PRC2.

Whether gene silencing and DNA hypermethylation could be maintained in the absence of PRC2 over-expression was also investigated as part of this chapter. These results indicated that *EZH2* and *SUZ12* knockdown in MSC 4 could relieve repression of PRC2 target genes with minimal promoter methylation - including *DKK2* and *TBX2* - although hypermethylated target genes (*EN1* and *RBP4*) remained silenced. Furthermore, methylation at the promoter regions of all PRC2 target genes was unaffected by *EZH2* or *SUZ12* knockdown. These results indicated that promoter hypermethylation and gene silencing could be maintained in the absence of PRC2 over-expression.

The relationship between PRC2 activity and DNA hypermethylation in cancer has been widely debated and remains the subject of current investigation. Two major models of this relationship have arisen from the findings in this field. On the one hand, it is hypothesized PRC2 activity marks target genes with

H3K27me3, leading to subsequent DNA hypermethylation and gene silencing (Vire et al., 2006). According to this model, reduced PRC2 activity – for example, by EZH2 knockdown – would lead to a concurrent reduction in the levels of H3K27me3 and DNA methylation at target gene promoters, resulting in re-activation of gene expression. The alternative model stipulates that H3K27me3 and DNA hypermethylation are independent mechanisms of gene silencing (Kondo et al., 2008). EZH2 knockdown would therefore only reverse silencing of genes repressed by H3K27me3 and not affect the expression of hypermethylated genes. The data presented in this chapter support the second model.

Several previous reports have noted a 'stem cell like' signature of hypermethylation at PRC2 target genes, suggesting that this aberrant DNA hypermethylation is directed by H3K27me3 early during transformation (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). These findings are supported by a study of EZH2 depletion and PRC2 target gene expression in U20S osteosarcoma cells (Vire et al., 2006). Following EZH2 knockdown, ChIP analysis of the *MYT1* and *WNT1* promoter regions showed reduced levels of H3K27me3, DNMT1 and DNA methylation at these loci. Additionally, these changes lead to increased expression of the previously silenced genes (Vire et al., 2006). These findings strongly support the notion that EZH2 activity and H3K27me3 can direct DNA methylation and subsequent gene silencing, as well as suggesting a direct interaction between EZH2 and DNMT1. In addition, these results suggest that continued *EZH2* expression is necessary to maintain gene silencing.

The requirement of PRC2 activity for the maintenance of target gene silencing in cancer has also been addressed through the use of dominant negative versions of histone H3 (Abbosh et al., 2006). Transgenic H3-K27R – containing a lysine to arginine mutation, preventing methylation at this residue – expression in CP70 ovarian cancer cells resulted in both genome-wide and gene specific hypomethylation alongside re-expression of genes previously silenced by H3K27me3 (Abbosh et al., 2006). However, another study suggested that DNA hypermethylation can remain in the absence of *EZH2* expression (McGarvey et al., 2007). Following *EZH2* knockdown in U20S and RKO cancer cells, the authors observed a genome-wide reduction in H3K27me3 levels. This knockdown had opposing effects on the expression of PRC2 target genes previously shown to display EZH2 dependent silencing in cancer (*hMLH, MYT, WNT1* and $p16^{INK4A}$). The results indicated that basally expressed genes containing minimal DNA methylation showed elevated expression following EZH2 knockdown. Conversely, densely hypermethylated promoters retained this methylation signature and remained silenced in the knockdown cells. These findings suggest that DNA methylation and resultant gene silencing can remain in the absence of *EZH2* over-expression in cancer cells (McGarvey et al., 2007).

The link between H3K27me3 and DNA hypermethylation in cancer has also been questioned by a study of chromatin modification and DNA methylation patterns in prostate cancer cells (Kondo et al., 2008). Analysis of genome wide H3K27me3 levels in PC3 cells by ChIP-chip identified 5% of promoters enriched for this modified histone. These promoters were specifically silenced in cancer and generally showed low promoter methylation levels whereas classical targets of DNA hypermethylation in cancer - such as *RASSF1* and *CDKN2A* -

showed no H3K27me3 enrichment. Furthermore, silencing of H3K27me3 target genes could be reversed by *EZH2* knockdown or treatment with the HDAC inhibitor TSA in the absence of DNA demethylation. Conversely, expression of genes displaying promoter hypermethyation could be activated by treatment with the demethylating agent 5-aza-deoxycytidine, but not following *EZH2* knockdown or TSA treatment. Finally, transfection of *RARβ2*, a gene normally silenced by H3K27me3, into PC3 and SW480 cells resulted in rapid and stable gene silencing associated with H3K27me3 but not DNA hypermethylation. The authors suggested that H3K27me3 mediated gene silencing is important for the repression of tumour suppressor genes in cancer, but that it is mechanistically distinct from gene silencing by DNA hypermethylation.

My analyses of H3K27me3 and DNA methylation levels in the MSC model are consistent with these findings, suggesting that the two mechanisms of gene silencing in cancer are independent. PRC2 target genes are enriched for H3K27me3 and show down-regulation of expression during step-wise transformation both in the presence and absence of promoter hypermethylation. Furthermore, target genes displaying hypermethylation do not retain this signature in MSC 5, despite continued silencing and H3K27me3 occupancy. In these cases, loss of promoter hypermethylation between MSC 4 and MSC 5 may be a result of the genome-wide hypomethylation observed in transformed MSC. Finally, *EZH2* knockdown relieved repression of unmethylated PRC2 target genes but did not affect DNA methylation levels at any of the promoter regions assayed. These findings suggest that PRC2 plays an important role in directing gene silencing from the earliest stages of step-wise transformation, but this silencing can take place without promoter DNA hypermethylation.

In the MSC model, PRC2 target genes show H3K27me3 enrichment from the earliest stage, in normal parental MSC. The observed decrease in H3K27me3 levels between these cells and transformed MSC 5 is contrary to previous comparisons of normal cells and their transformed equivalents. For example, a comparison of H3K27me3 levels between normal peripheral blood cells and cultured prostate epithelial cells to prostate cancer cells only noted H3K27me3 enrichment and gene silencing in the cancer cells (Kondo et al., 2008). Differences in the epigenetic profile of normal stem cells and differentiated cells may therefore affect the role and requirement of PRC2 up-regulation during transformation. In other words, stem cells may not require PRC2 up-regulation for transformation to take place because they already possess a permissive 'stem cell like' epigenetic state.

Several previous studies have shown elevated expression of PRC2 components in a variety of cancers and shown that *EZH2*, the key catalytic component of the complex, has properties characteristic of an oncogene (Sparmann and van Lohuizen, 2006). EZH2 depletion by siRNA in prostate cancer and multiple myeloma cells has been shown to inhibit cell proliferation *in vitro* (Croonquist and Van, 2005; Varambally et al., 2002). Furthermore, *EZH2* over-expression induces growth factor independence in multiple myeloma as well as anchorage independent growth and invasion of immortalized human mammary epithelial cells (Croonquist and Van, 2005; Kleer et al., 2003). Although this suggests that PRC2 components play an important role in malignant cells, it is not known whether over-expression of this complex is required during the transformation process. My results indicate that knockdown of *EZH2* and *SUZ12* in MSC 4 cells does not prevent subsequent transformation

of these cells with the addition of H-Ras^{V12}, suggesting that transformation of MSC can take place in the absence of PRC2 over-expression and that *EZH2* does not necessarily act as an oncogene in this model.

In conclusion, this study is the first to investigate changes in PRC2 expression and activity during step-wise transformation. My results showed a significant and specific increase in *EZH2*, *SUZ12* and *EED* mRNA expression between parental and transformed MSC, with EZH2 and SUZ12 protein levels also increasing. GSEA indicated a functional output of this over-expression, with down-regulation of PRC2 target gene expression between MSC 0 and MSC 4. ChIP analysis of selected target genes indicated that H3K27me3 is present at these genes in MSC 0, but that levels decreased during step-wise transformation. Methylation analysis of these genes demonstrated that although some show hypermethylation during transformation, none retain this methylation signature in MSC 5, contrary to previous reports of PRC2 target gene hypermethylation in cancer. Finally, stable knockdown of *EZH2* in MSC 4 reversed repression of unmethylated PRC2 target genes, though hypermethylated target genes remained unexpressed. Furthermore, the methylation status of target gene promoters did not change following *EZH2* knockdown in MSC 4, suggesting that PRC2 mediated gene repression is independent of promoter hypermethylation. Finally, both *EZH2* and *SUZ12* knockdown cells could be transformed following over-expression of oncogenic H-Ras^{V12}, suggesting that PRC2 does not function as an oncogene in this model. These findings suggest that PRC2 plays an important early role in transformation and that target gene silencing by H3K27me3 and DNA hypermethylation may be independent events during transformation.

Chapter 6: Conclusions

Hallmark epigenetic changes, including genome-wide hypomethylation, gene specific hypermethylation and characteristic chromatin modifications, have been well documented in cancer (Esteller, 2006; Feinberg and Tycko, 2004; Fraga et al., 2005; Herman and Baylin, 2003; Jones and Baylin, 2002). Although these changes have been noted in many comparisons of normal and malignant cells, relatively little is known about the importance or timing of epigenetic changes *during* the transformation process. The overall aim of this thesis was to investigate these epigenetic changes in a genetic step-wise model of transformation in mesenchymal stem cells. This study is the first to show changes in genome-wide methylation levels during step-wise transformation and to show that this hypomethylation is induced by the expression of oncogenic H-Ras^{V12}. Although hypomethylation was observed in the MSC model, this change was not required for transformation to take place. As well as genome-wide hypomethylation, other hallmark epigenetic changes including gene specific hypermethylation and characteristic histone modifications have been shown to occur in transformed cells. In order to investigate the mechanisms governing these changes in the MSC model, the role of the epigenetic silencing complex PRC2 during step-wise transformation was investigated. My study has shown that up-regulation of PRC2 expression takes place during step-wise transformation along with repression of PRC2 target genes. Down-regulation of target gene expression can take place both with or without DNA

hypermethylation, suggesting that these silencing mechanisms may act independently.

6.1 Relevance of the MSC Model

Many studies of cancer consider differences between normal and fully transformed tissues and are therefore unable to distinguish between early and late changes during transformation or determine the relationships between these changes. The MSC model was originally developed to study the events occurring during transformation in a controlled manner. As well as including a series of genetic hits reflecting typical genetic aberrations that occur during neoplastic development, this model also displays several characteristics of *in vivo* transformation including up-regulation of anti-apoptotic pathways, up-regulation of angiogenic genes and down-regulation of MHC class I and II molecules (Funes et al., 2007). Furthermore, this thesis has demonstrated that hallmark epigenetic changes commonly observed in cancer also take place in this model, suggesting that it provides a relevant and accurate representation of transformation processes.

Although *in vitro* models of cancer are widely used in research, it is important to recognize the differences between these models and *in vivo* transformation processes. First, *in vitro* cancer cell line cultures have proved valuable tools for the study of neoplastic development since these cells can be easily grown and manipulated, allowing numerous experiments to be conducted that would not otherwise be possible with primary tissue samples. However, *in vitro* cell culture conditions do not accurately mimic the *in vivo* micro-

environment, and may therefore affect the growth and behaviour of cultured cancer cells (van Staveren et al., 2009). Furthermore, it is possible that continued *in vitro* culture of cells may introduce epigenetic changes that would not arise during *in vivo* growth. For example, an analysis of genome-wide DNA methylation patterns in mesenchymal stromal cells demonstrated that although overall methylation patterns were maintained during long term cell culture, highly significant differences were observed at specific CpG sites, similar to those observed during *in vivo* aging (Bork et al., 2010). This is an important consideration when interpreting results obtained from the step-wise model of transformation in MSC, since later stages in the model have been cultured for a much greater length of time relative to parental MSC. However, the stability of cellular phenotypes in the MSC model and reproducibility of results obtained in this thesis suggest that these observations are a result of the genetic hits introduced in this model rather than epigenetic changes arising from continued cell culture.

Furthermore, the link between many epigenetic regulatory mechanisms and cell proliferation must be considered when interpreting results from this thesis in light of the difference in proliferative rate between the MSC cell lines. In particular, cell cycle length will determine the proportion of cells in S-phase at any given time, therefore influencing genome-wide methylation levels in an unsynchronized population of these cells. Additionally, changes in the expression of E2F responsive may account for some of the observed changes in gene expression during step-wise transformation, as demonstrated by differences in gene expression microarray analysis when corrected for expression of the proliferation marker PCNA. The complex interaction between cell cycle length,

gene expression and DNA methylation levels must therefore be carefully considered when interpreting findings from this model and in the design of future experiments.

Although results from this thesis demonstrate that genetic aberrations can precede and induce epigenetic changes, this may not reflect the order in which these events occur during *in vivo* transformation. Several previous models have described the tumourigenic process through a series of genetic and epigenetic aberrations (Baylin and Ohm, 2006; Fearon and Vogelstein, 1990; Feinberg et al., 2006). Although the MSC model is similar to these in that it involves a sequential series of events necessary and sufficient to induce transformation, this model differs in the nature and order of these changes. Both the Baylin and Vogelstein models include epigenetic changes such as genome-wide hypomethylation as early events in the step-wise transformation process. Conversely, the MSC model is based on a series of genetic events that induce significant epigenetic changes such as genome-wide hypomethylation and selected PRC2 target gene silencing in the later stages of transformation. It remains possible that these differences are reflective of the varying transformation requirements in the development of carcinomas and sarcomas.

Further differences between *in vitro* and *in vivo* transformation processes may arise from the choice of an adult stem cell as the starting point for this model. The cancer stem cell hypothesis stipulates that tumour-initiating cells share many properties with normal stem cells (Reya et al., 2001; Visvader and Lindeman, 2008). It is believed that cancer stem cells arise from transformation events affecting normal stem cells, or that fully differentiated cells may 'de-differentiate' towards a more stem cell-like state through epigenetic alterations

(Zilberman, 2007). By utilizing MSC as the starting point for step-wise transformation, this model may not accurately reflect the importance of early epigenetic 'de-differentiation' processes in neoplastic development. However, the cellular origin of cancer stem cells remains a subject of investigation and it is possible that transformation of both stem cells and differentiated cells takes place *in vivo*.

Despite these considerations, this step-wise model of transformation remains a valuable and relevant tool for the investigation of changes occurring during the development of cancer. The validity of this model has been confirmed by extensive characterization of all MSC cell lines, demonstrating expected downstream changes in gene expression on the introduction of each genetic hit. Furthermore, tumours formed by growth of MSC 5 cells in nude mice show gene expression patterns that cluster with the expression profiles of primary human spindle cell sarcoma, indicating that the *in vitro* transformed MSC are very similar to *in vivo* transformed from the same lineage (Funes et al., 2007). Furthermore, this model is particularly powerful for studies of the transformation process since a completely matched normal cell control can be compared to the transformed equivalent. Results obtained from this model therefore have direct relevance to cancer and can offer novel insight into carcinogenesis. This thesis demonstrates that epigenetic changes occur during step-wise transformation, and that there is a complex interplay between genetic and epigenetic changes during transformation. Furthermore, it suggests that hallmark epigenetic changes observed *in vitro* can arise from the step-wise genetic hits utilized in the MSC model.

6.2 Genome-wide Hypomethylation

Although genome-wide hypomethylation is one of the most widely documented epigenetic changes in cancer, little is known about the cause of this loss of methylation. In this study, I aimed to investigate whether this epigenetic change occurs during step-wise transformation of MSC, and to identify the step at which this loss of methylation takes place.

The correlation between repetitive element methylation and genome-wide methylation levels has been well documented, with these sequences often displaying hypomethylation in cancer (Ehrlich, 2002b; Weisenberger et al., 2005). Analysis of repetitive element methylation using the qPCR based method MethyLight was therefore conducted to provide a measure of genome-wide methylation in the MSC model. This method was chosen as it is high throughput, well optimized and can be performed with small amounts of DNA. Comparison of Alu, Sat2 and LINE1 methylation levels as measured by MethyLight between parental and transformed MSC and HF revealed a significant decline in Sat2 and LINE1 methylation in MSC 5, occurring at the introduction of H-Ras in the final oncogenic step. Hypomethylation of these elements was not observed in a parallel step-wise model of transformation in HF. The MethyLight data were confirmed by bisulphite sequencing, indicating significant loss of methylation at specific sites within the Sat2 element. This study is the first to definitively address the timing of genome-wide hypomethylation during transformation and to demonstrate repetitive element hypomethylation following H-Ras^{V12} expression.

Although DNA hypomethylation in cancer has been shown to affect a wide variety of sequences throughout the genome, it is believed that much of the observed hypomethylation occurs at repetitive elements, which together account for ~45% of the genome (Ehrlich, 2002b; Gama-Sosa et al., 1983c; Weisenberger et al., 2005). Results obtained from the MSC model are consistent with these findings, indicating a significant loss of methylation at the Sat2 and LINE1 repetitive elements in MSC 5, suggesting that transformed MSC display genome-wide hypomethylation. Although I attempted to measure genome-wide methylation levels in these cells with a novel FACS based approach, this method did not indicate a significant difference in methylation between MSC 4 and MSC 5. Further investigation by nuclease digestion of total nuclear DNA followed by HPLC analysis is required to quantitatively determine whether genome-wide methylation levels are significantly lower in MSC 5 compared to the other MSC cell lines. Nonetheless, the FACS method remains a valuable tool for future studies of genome-wide hypomethylation, and provides an attractive alternative to HPLC for quantitative analysis of genomic methylcytosine content. This method has several advantages over the available alternatives in that it can be used on a very small number of cells and provides methylation data at a single cell resolution. Additionally, the method could be used in conjunction with co-staining for other markers of interest to investigate correlations between marker expression and genome-wide methylation levels.

A number of previous studies have linked Ras signaling to changes in DNA methylation. Ras mediated hypermethylation of FAS, MMP2, DNMT1 and clusterin has been observed in a variety of transformed cells (Gazin et al., 2007; Lund et al., 2006; MacLeod et al., 1995). Furthermore, v-Ha-Ras overexpression

has been demonstrated to induce a demethylase activity in mouse embryonal P19 cells, leading to demethylation of endogenous and exogenous methylated sequences (Szyf et al., 1995). My findings are consistent with these studies, indicating that H-Ras signaling plays an important role in the regulation of DNA methylation patterning, and can induce genome-wide hypomethylation in transformed human cells.

This investigation of genome-wide methylation has also provided insight into the timing of this epigenetic change during transformation. Genome-wide hypomethylation has traditionally been viewed as an early event in carcinogenesis, and has been included as such in many well established models of transformation (Fearon and Vogelstein, 1990; Feinberg et al., 2006). This view is supported by the observation of genome-wide hypomethylation in benign neoplasms and normal tissues adjacent to tumours (Feinberg et al., 2006; Goelz et al., 1985; Suter et al., 2004). My results suggest the opposite, indicating that genome-wide hypomethylation can occur at the final stages during step-wise transformation. This observation concurs with several other studies which demonstrate genome-wide hypomethylation as a late event during cancer progression (Ehrlich et al., 2006; Gama-Sosa et al., 1983a; Yegnasubramanian et al., 2008). However, by using a series of genetic hits to model events during transformation, this system may not accurately reflect the importance of early epigenetic changes such as hypomethylation during *in vivo* transformation.

The relationship between genetic and epigenetic aberrations during transformation has also been widely studied. Both genetic and epigenetic changes can lead to the activation of oncogenes and silencing of tumour suppressor genes, acting concurrently to contribute towards transformation

(Esteller, 2006; Feinberg and Tycko, 2004; Jones and Baylin, 2002). For example, simultaneous genetic and epigenetic disruptions to DNA repair and Wnt signaling pathways contribute towards the transformed phenotype of HCT116 colon cancer cells (Herman et al., 1998; Morin et al., 1997; Suzuki et al., 2004). However, the order in which these two types of disruption take place remains unclear. It has been suggested that early epigenetic changes may lead to subsequent genetic abnormalities in one of two ways – either through epigenetic disruption of pathways involved in DNA repair leading to an elevated mutation rate, or through early oncogene addiction that is subsequently maintained through genetic changes (Baylin and Ohm, 2006). In support of this first hypothesis, epigenetic inactivation of MGMT and hMLH1 can lead to the accumulation of specific mutations through ineffective DNA repair, and silencing of CDKN2A may allow breast epithelial cells to escape senescence and accumulate genetic abnormalities (Esteller et al., 1999; Kane et al., 1997; Kiyono et al., 1998; Romanov et al., 2001).

The second hypothesis is based on oncogene addiction – the notion that transformed cells become dependent on abnormal oncogene products or activation of specific pathways (Weinstein, 2002). Although originally applied to aberrant genetic changes in cancer, it is becoming clear that early epigenetic disruptions may also lead pre-malignant cells to become dependent on dysfunctional pathways. This could predispose these cells to additional genetic alterations in the same pathways, further increasing the dependence of such cells on these abnormalities. For example, aberrant activation of the Wnt signaling pathway through APC mutation is known to play an important role in the early progression of colorectal cancer, with restoration of APC function in cultured

colon cancer cells resulting in apoptosis (Gregorieff and Clevers, 2005; Weinstein, 2002). This signaling pathway can also be activated during early stages of the transformation process through epigenetic changes in the absence of genetic mutations. The promoters of *SFRP* gene family members, which normally act to antagonize Wnt signaling, have been shown to be aberrantly methylated and silenced in pre-invasive colon lesions (Suzuki et al., 2004; Taketo, 2004). Epigenetic activation of the Wnt signaling pathway is essential for continued growth of HCT116 colon cancer cells, with *DNMT* deletion causing loss of methylation and re-activation of *SFRP* expression, leading to reduced Wnt signaling and subsequent cellular apoptosis (Baylin and Ohm, 2006). Epigenetic mechanisms may therefore play an important role in the early expansion of pre-malignant cells, which accumulate further genetic and epigenetic changes leading to transformation.

Although these examples demonstrate that early epigenetic disruptions can lead to oncogene addiction and subsequent genetic changes in similar pathways, my results suggest the opposite. In the MSC model, genetic hits lead to subsequent epigenetic changes, which together contribute towards the transformation process. This suggests that genetic changes can precede and induce epigenetic changes, and that the order of these events may be flexible for transformation to occur. However, as noted earlier, the nature of our MSC model may not accurately reflect the order of events during *in vivo* transformation, where epigenetic changes may play a more important role in the early expansion of cells.

Further characterization of the MSC cell lines is necessary to determine the consequences of repetitive element hypomethylation in this model. Aberrant

transcription of repetitive elements, one potential effect of hypomethylation, could be investigated by qRT-PCR (Kochanek et al., 1995; Yoder et al., 1997). Similarly, active transposition of LINE1 elements may occur following hypomethylation, and could be monitored by analysis of repetitive element copy number by quantitative PCR (Coufal et al., 2009). Finally, repetitive element hypomethylation may lead to increased recombination events between these areas of high sequence similarity, causing karyotypic instability (Chen et al., 1998; Chen et al., 2007; Hsieh and Lieber, 1992). For example, characteristic karyotypic changes following Sat2 hypomethylation have been observed in ICF syndrome (Ehrlich, 2002a; Tuck-Muller et al., 2000). Although my results show severe karyotypic abnormalities in the MSC 5 population, without further karyotypic analysis of the other MSC cell lines, it is not possible to establish whether the chromosomal changes observed in MSC 5 are the result of DNA hypomethylation. Additional karyotypic analysis following the time course of H-Ras induced hypomethylation will also be required to test whether these changes are caused by loss of repetitive element methylation during step-wise transformation of MSC.

This study is the first to investigate epigenetic changes during step-wise transformation of MSC with a defined series of genetic hits. My results indicate that repetitive element hypomethylation can be induced in this system, and that this loss of methylation occurs after the expression of H-Ras^{V12} in the final oncogenic hit. This indicates that in the MSC model, genome-wide hypomethylation is a late event during transformation and demonstrates that genetic changes can precede and induce epigenetic changes.

6.3 *Conditional Expression of ER^{TM}-H-Ras^{V12}*

To gain further insight into the potential mechanism of hypomethylation, I aimed to investigate the time course of this loss of methylation following H-Ras^{V12} expression. A conditional H-Ras^{V12} system, based on a fusion protein comprising H-Ras^{V12} and a mutated estrogen receptor domain, ER^{TM}, was utilized for this investigation. The system was extensively characterized and shown to transform cells when H-Ras^{V12} expression was induced with 4-OHT. However, analysis of repetitive element methylation following induced H-Ras^{V12} expression failed to detect hypomethylation when cells were grown in the presence of 4-OHT for one month.

Expression of constitutive H-Ras^{V12} in MSC 4 was repeated in order to determine whether repetitive element hypomethylation could be recapitulated in this model. Sat2 methylation was assessed by MethyLight in these cells one week post infection and following growth of the cell lines for one month. These data showed no change in repetitive element methylation after one week, but demonstrated Sat2 hypomethylation in H-Ras^{V12} expressing cells one month after retroviral infection. This indicates that loss of repetitive element methylation following H-Ras^{V12} expression takes place gradually, suggesting a passive mechanism of hypomethylation.

The widespread incidence of genome wide hypomethylation in a range of cancers, as well as in benign neoplasms, has suggested that this epigenetic change occurs early during the process of transformation (Feinberg et al., 2006). The results obtained from this step-wise model of transformation, however, indicate that hypomethylation occurred only in fully transformed cells following

the final oncogenic 'hit'. Furthermore, the loss of methylation was only observed after 4 weeks in culture following transformation. This suggests that hypomethylation occurred as a late event in step-wise transformation via slow passive loss of methylation. However, transformation in this model is induced by a series of genetic hits and may not fully recapitulate the true early epigenetic changes that occur during *in vivo* transformation as suggested by the epigenetic progenitor origin of cancer model (Feinberg et al., 2006). It is possible, therefore, that although genome-wide hypomethylation is a late event in this model, it may occur earlier during *in vivo* cancer progression as a prelude for further genetic and epigenetic changes.

Using conditional expression of the fusion protein ER^{TM}-H-Ras, I was unable to recapitulate the loss of methylation observed with the constitutive expression of H-Ras. Despite the lack of hypomethylation, these cells were transformed as shown by their ability to form colonies in soft agarose. This suggested that genomic hypomethylation was not essential for *in vitro* transformation of these cells. The potential causal link between genome-wide hypomethylation and transformation has been investigated in a number of previous studies, with conflicting findings. Genome-wide loss of methylation in mice containing a hypomorphic allele of *dnmt1* can result in transformation and increased incidence of soft tissue sarcomas, possibly through increased incidence of karyotypic changes and chromosomal abnormalities (Eden et al., 2003). However, another mouse model of genome wide hypomethylation induced by *dnmt1* knockdown notes opposing effects on the growth of different tumour types (Yamada et al., 2005). Together, these reports show that genome wide hypomethylation can promote some instances of transformation but there is no

definitive demonstration that hypomethylation is a strict requirement for transformation.

The timing of genome wide hypomethylation during the development of cancer could also provide an indication of whether this change is causative of transformation. Many studies have suggested that global DNA hypomethylation is an early event in transformation, and can be observed in benign tumours or adjacent normal cells (Feinberg et al., 1988; Goelz et al., 1985; Suter et al., 2004). Global hypomethylation is also included as an event in various models of step-wise transformation in colorectal cancer (Fearon and Vogelstein, 1990; Feinberg et al., 2006). The consensus from these studies is that global hypomethylation in the adjacent normal tissue indicates this change is an early event in the transformation process, potentially leading to karyotypic instability and further oncogenic changes. However, several other investigations have failed to detect reduced global methylcytosine levels in benign neoplasms (Ehrlich et al., 2006; Gama-Sosa et al., 1983b). Indeed, many studies have also failed to observe early hypomethylation during malignant tumor progression (Florl et al., 2004; Yegnasubramanian et al., 2008). My study agrees with the latter observation, suggesting that genomic hypomethylation is not an early oncogenic change and thus is unlikely to be a cause of transformation.

The reason as to why the conditional ER^{TM}-H-Ras did not promote hypomethylation, when the constitutive H-Ras did, is not clear. Differences in the rate of methylation loss between the two vector systems may account for this effect. Alternatively, although the activity of the conditional ER^{TM}-H-Ras^{V12} was extensively characterized, differences between this fusion protein and H-Ras^{V12} could also have affected the process of hypomethylation. Additional

characterization of these two systems by gene expression microarray analysis may provide further insight into functional differences between the two versions of H-Ras and the mechanism of genomic hypomethylation during transformation.

Analysis of gene expression microarray data for a number of epigenetic regulators and metabolic enzymes involved in DNA methylation indicated numerous alterations to the expression of these genes but failed to conclusively indicate the cause of hypomethylation in MSC 5. Furthermore, different patterns gene expression are evident when the GEM data are corrected for cell proliferation rate in each MSC cell line. The effect of cell proliferation on expression of epigenetic regulators should therefore be taken into account when designing further experiments in this model, with appropriate correction factored into the analysis of future data.

The microarray data and qRT-PCR analysis of mRNA levels in the MSC cell lines did demonstrate up-regulation of GADD45A, a putative DNA demethylase, between MSC 4 and MSC 5. Finally, this analysis also suggested that the metabolic pathways surrounding S-adenosylmethionine (SAM) production were deregulated in MSC 5. Together, these data indicate that during step-wise transformation, there are changes to the expression of enzymes involved in the regulation of DNA methylation, potentially leading to disruption of normal DNA methylation processes. However, no change in expression that would individually account for genome-wide hypomethylation in MSC 5 is evident.

Both elevated and reduced DNMT expression has been noted in previous studies of cancer and so far no clear link between altered expression of these enzymes and incidence of genome-wide hypomethylation has emerged (Ehrlich et al., 2006; Li et al., 2003). Although *DNMT1* mRNA expression does not increase during step-wise transformation when GEM data are corrected for cell proliferation, analysis of methyltransferase activity in the MSC cell lines will be necessary to determine whether the activity of this class of enzymes is altered during transformation. Similarly, the up-regulation of GADD45A expression at this stage suggests an attractive hypothesis of active demethylation taking place in MSC 5. However, the involvement of this enzyme in demethylation remains controversial, and a link between increased GADD45A expression and genome-wide hypomethylation in cancer has not been previously shown (Barreto et al., 2007; Jin et al., 2008; Rai et al., 2008). This could be tested through GADD45A knockdown studies in MSC 4 or MSC 5 followed by analysis of Sat2 methylation levels. *In vitro* demethylation assays will also be necessary to show that this enzyme promotes active demethylation in human cells.

Deregulated methionine metabolism and S-adenosylmethionine (SAM) biosynthesis has also been suggested to play an important role in transformation (Laird and Jaenisch, 1996; Szyf et al., 2004). Analysis of SAM levels during step-wise transformation would provide an indication of whether levels of this key substrate are limiting in MSC 5 (Wagner et al., 1984). Manipulation of the biochemical pathways surrounding SAM biosynthesis, by siRNA or drug treatment, would also be valuable in determining whether disruptions to these pathways can cause repetitive element hypomethylation in MSC .

The mechanism by which H-Ras^{V12} signaling induces these changes will also be an interesting subject of investigation. Downstream Ras signaling can occur via the Raf, PI3K and Ral-GEF pathways, with activation of different combinations of these pathways required for transformation in different cell types (Rangarajan et al., 2004). The involvement of a particular pathway in repetitive element hypomethylation could be studied by introducing mutant versions of Ras, capable of only signaling through one downstream pathway, in different combinations into MSC 4 (Rangarajan et al., 2004). Alternatively, drug treatment could be used to block individual pathways in MSC 5. Identification of the downstream signaling pathways responsible for the induction of repetitive element hypomethylation will provide valuable insight into the mechanism of this loss of methylation. Additionally, differences in downstream signaling between the constitutive H-Ras^{V12} and conditional ER^{TM}-H-Ras^{V12} may be responsible for the lack of hypomethylation when MSC are transformed using the conditional system.

In summary, although genome wide hypomethylation has traditionally been viewed as an early event in oncogenesis, there is ample evidence supporting both this view and the alternative, that hypomethylation can occur later in transformation or not at all. Here, I show that hypomethylation can occur late during step-wise transformation, although it was not a necessary event for transformation to take place. Furthermore, in the instance where hypomethylation was observed, this was a late event in this step-wise genetic model, occurring subsequent to a common genetic aberration in cancer. In conclusion, this study has demonstrated a link between H-Ras and hypomethylation of repetitive elements in a stepwise model of transformation in

human MSC. This model provides a controlled background in which known genetic hits can be linked to observed epigenetic changes. My results provide a further step towards elucidating a mechanistic basis for genome wide hypomethylation in cancer and a valuable model for establishing the consequences of this change.

6.4 The Role of PRC2 During Step-wise Transformation of MSC

The final results chapter of this thesis aimed to address other hallmark epigenetic changes in cancer, namely gene specific hypermethylation and chromatin modifications. The expression of PRC2, an important epigenetic regulator complex, is known to be deregulated in a variety of cancers, leading to aberrant gene silencing. Gene expression microarray analysis indicated that all three core components of the PRC2 complex were significantly up-regulated in MSC 5 compared to MSC 0. This suggested that deregulated activity of this complex could be an important event during step-wise transformation. I aimed to confirm this up-regulation and address potential downstream effects on PRC2 target gene expression, methylation and chromatin modification.

Expression of *EZH2*, *SUZ12* and *EED* mRNA and protein were assessed by qRT-PCR and western blot respectively in all MSC cell lines, confirming the up-regulation observed in the microarray analysis. These findings are consistent with numerous other studies of cancer, which show elevated expression of PRC2 components in transformed cells (Sparmann and van Lohuizen, 2006).

Furthermore, the MSC model provides a valuable tool for further investigation of the mechanisms by which PRC2 up-regulation occurs during transformation. Although several pathways, including E2F and MYC, have been shown to regulate *EZH2*, *SUZ12* and *EED* expression, many of the genetic aberrations employed in this step-wise model of transformation also appear to cause up-regulation of these genes (Bracken et al., 2003; Weinmann et al., 2001). For example, data from this model suggest that hTERT and HPV E6 expression (abrogating the activity of p53) could cause up-regulation of EZH2. Further investigation of these observations could uncover novel regulatory mechanisms for PRC2 component expression.

Global levels of H3K27me3 in all MSC cell lines were also analysed by western blot, indicating a reduction in the relative levels of this modified histone in MSC 1, MSC 3, MSC 4 and MSC 5 compared to parental MSC. One possible explanation of this observation could be altered substrate specificity of EZH2 when associated with different EED isoforms in transformed cells (Kuzmichev et al., 2005). A shift in the relative levels of EED isoforms during step-wise transformation may affect PRC2 specificity or activity, leading to a reduction in genome-wide levels of H3K27me3 in favour of H1K26 methylation. However, it is not possible to verify this hypothesis without further characterization of EED isoform levels and the substrate specificity of EZH2 containing complexes in this system.

PRC2 target gene expression during step-wise transformation was also investigated by GSEA. Gene expression microarray data for a comprehensive list of PRC2 target genes identified in ES cells were analysed, revealing significant down-regulation of this gene set in MSC 4 compared to MSC 0 and between

MSC 4 compared to MSC 3 (Boyer et al., 2006). Gene expression microarray data for GSEA core genes confirmed this trend, as did qRT-PCR analysis of PRC2 target gene mRNA levels. These data suggest that PRC2 up-regulation has a functional effect in the MSC model, leading to repression of target genes during step-wise transformation. This is in line with previous reports of PRC2 target gene repression in cancer (Squazzo et al., 2006).

PRC2 mediated down-regulation of target gene expression results from creation of the repressive H3K27me3 histone modification at gene promoters, and may also be caused by direct interactions between this complex and DNMT1, leading to promoter DNA hypermethylation (Kirmizis et al., 2004; Sparmann and van Lohuizen, 2006). Selective hypermethylation of PRC2 target genes in cancer has been noted in a number of studies, suggesting that PRC2 up-regulation during transformation may contribute to the transformed phenotype through aberrant silencing of these genes (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). This hypothesis is supported by the identification of direct interactions between EZH2 and DNMT1 in an investigation of osteosarcoma cells (Viré et al., 2006). However, other studies of *EZH2* knockdown have questioned the link between H3K27me3 and PRC2 target gene hypermethylation in cancer, showing that DNA methylation and gene expression levels do not always change when levels of this key epigenetic regulator are reduced (Kondo et al., 2008; McGarvey et al., 2007).

In order to investigate the mechanism by which PRC2 target genes are down-regulated in the MSC model, levels of H3K27me3 and DNA methylation at selected PRC2 target genes were analysed by chromatin immunoprecipitation (ChIP). This demonstrated that PRC2 target gene promoters were enriched for

H3K27me3 compared to a non-PRC2 target gene control. However, levels of this histone modification declined significantly between MSC 0 and MSC 5 at the regions assayed. Investigation of DNA methylation levels at four PRC2 target genes also produced a number of interesting results. Although all these genes showed significantly reduced expression between MSC 0 and MSC4, only two displayed any degree of hypermethylation during step-wise transformation. Furthermore, none of the genes assayed were hypermethylated in MSC 5. This suggests that silencing of PRC2 target genes during step-wise transformation can take place both with and without DNA hypermethylation.

To establish whether this gene silencing could be maintained in the absence of PRC2 over-expression, *EZH2* and *SUZ12* were knocked-down in MSC 4. This cell line was chosen for the experiment since the most significant down-regulation of PRC2 target genes was observed at this stage of step-wise transformation. Analysis of target gene expression in these knock-down cells indicated up-regulation of *DKK2* and *TBX2* following *EZH2* knockdown. However, DNA methylation levels at the PRC2 target genes assayed did not change significantly following knockdown. Together, down-regulation of PRC2 target gene expression during step-wise transformation both with and without DNA hypermethylation alongside reversal of this silencing in the absence of any changes in promoter methylation suggest that H3K27me3 and DNA methylation are independent gene silencing mechanisms, contrary to the hypothesis that H3K27me3 directs DNA methylation (Viré et al., 2006).

Finally, the transformation potential of MSC 4 following *EZH2* or *SUZ12* knock-down was addressed by overexpression of H-Ras^{V12} in these cell lines. Both *EZH2* and *SUZ12* knockdown cells were transformed and produced

GFP positive colonies in soft agarose, indicating continued expression of the shRNA constructs. These data showed that transformation of MSC can take place in the absence of *EZH2* or *SUZ12* over-expression and that these genes do not necessarily act as typical oncogenes in this model. These results are contrary to previous reports of *EZH2* acting as an oncogene in prostate and breast cancer cells (Croonquist and Van, 2005; Kleer et al., 2003; Varambally et al., 2002). Maintenance of PRC2 target gene silencing by continued DNA hypermethylation following *EZH2* knockdown in MSC 4 may account for the difference in oncogenic potential of *EZH2* in this model. Additionally, the genetic hits utilized in this step-wise model of transformation may contribute towards silencing of PRC2 target genes through alternative mechanisms, removing the requirement of epigenetic gene repression by PRC2 for transformation to take place.

Future work will involve further characterization of histone modifications at PRC2 target gene promoters by ChIP. Analysis of gene specific H3K27me3 signatures in all MSC, and PRC2 knockdown cell lines would be valuable in confirming the continued presence of this mark throughout step-wise transformation. Additionally, the relationship between this histone modification and DNA methylation in this model could be further explored by treating cells with demethylating agents or inhibitors of HDAC activity and observing the consequences of these interventions on expression of PRC2 target genes. Finally, although the candidate gene approach adopted in this study has produced several interesting results, a comprehensive picture of genome-wide changes to H3K27me3 patterning during step-wise transformation can only be obtained by ChIP followed by microarray hybridization (ChIP-chip) or high throughput sequencing (ChIP-seq).

Down-regulation of PRC2 target gene expression was observed during step-wise transformation despite reduced levels of H3K27me3 at gene promoters. Other histone modifications may therefore play a role in the regulation of PRC2 target gene expression and incidence of DNA hypermethylation in this model. H3K27me3 is one part, along with H3K4me3, of the 'bivalent' chromatin state present at many key developmental genes in stem cells (Bernstein et al., 2006). PRC2 target gene down-regulation may therefore result from loss of the activating H3K4me3 mark rather than an increase in H3K27me3 enrichment. On this basis, analysis of H3K4me3 levels at PRC2 target genes may provide additional insight into the mechanisms by which PRC2 target genes are silenced during step-wise transformation, and how silencing of some genes can remain following *EZH2* or *SUZ12* knockdown.

Several studies have also noted the importance of H3K9 acetylation and methylation for gene repression in cancer and the association between this histone modification and DNA methylation (Bachman et al., 2003; Kondo et al., 2004). It is possible that the patterns of DNA methylation observed at PRC2 target genes may also be directed by this histone modification. Further investigation of H3K9 acetylation and other histone modifications in the MSC model is essential in gaining a more comprehensive understanding of epigenetic gene regulation during step-wise transformation.

Further analysis of DNA methylation at PRC2 target genes would also be useful in elucidating the interaction between chromatin modifications and DNA methylation. Although a candidate gene approach was adopted in this study, the most useful insight into PRC2 target gene methylation would be profiling of all MSC cell lines in a genome-wide manner by methylated DNA

immunoprecipitation (MeDIP) (Weber et al., 2005). This would provide an unparalleled overview of gene specific changes to DNA methylation during step-wise transformation, and would allow correlations between DNA methylation and alterations in gene expression to be studied in more detail.

Overall, this thesis has demonstrated that hallmark epigenetic changes occur during step-wise transformation of MSC, suggesting a model in which both genetic and epigenetic changes contribute towards transformation (Figure 6.1). Analysis of genome-wide methylation levels by MethyLight indicated that hypomethylation of repetitive element occurs on expression of oncogenic H-Ras^{V12} in MSC 5. Up-regulation of PRC2 also takes place during transformation, leading to target gene silencing both with and without gene specific hypermethylation. These studies show that cancer specific epigenetic changes occur during step-wise transformation and suggest that tumour-associated epigenetic changes can follow the genetic aberrations introduced in this model. This model is valuable and relevant to further explore the mechanisms behind epigenetic alterations in cancer.

Figure 6.1

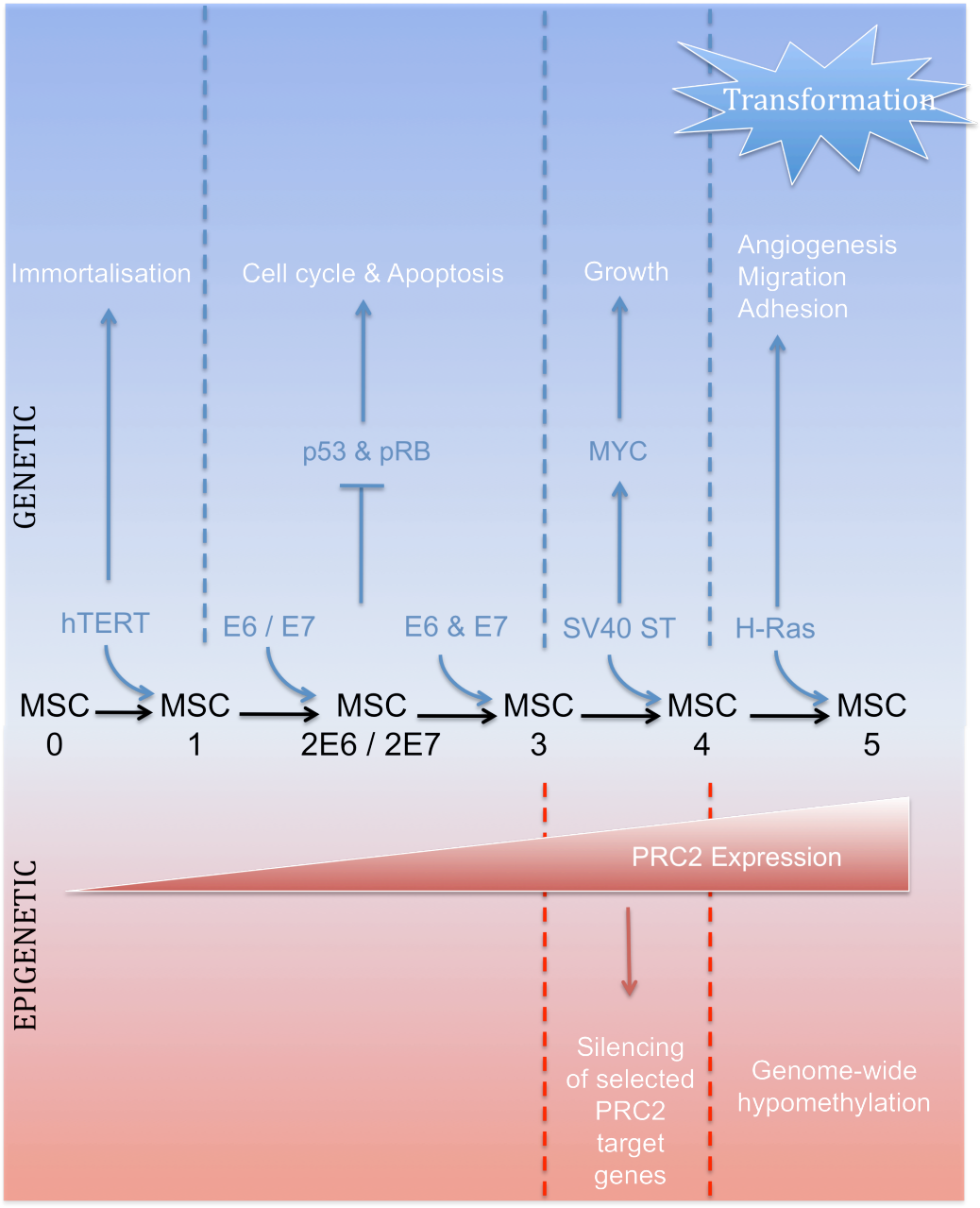


Figure 6.1: Hallmark epigenetic changes occur during step-wise transformation of MSC. MSC were transformed through the sequential addition of five genetic 'hits' (hTERT, HPV E6 and E7, SV40 Small-T and H-Ras^{V12}). Studies of this model conducted as part of this thesis have indicated that hallmark epigenetic changes also take place during step-wise transformation, including genome-wide hypomethylation, increased PRC2 expression and specific silencing of PRC2 target genes.

References

Abbosh, P.H., Montgomery, J.S., Starkey, J.A., Novotny, M., Zuhowski, E.G., Egorin, M.J., Moseman, A.P., Golas, A., Brannon, K.M., Balch, C.*, et al.* (2006). Dominant-negative histone H3 lysine 27 mutant derepresses silenced tumor suppressor genes and reverses the drug-resistant phenotype in cancer cells. Cancer Res *66*, 5582-5591.

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci USA *100*, 3983-3988.

Austin, T.W., Solar, G.P., Ziegler, F.C., Liem, L., and Matthews, W. (1997). A role for the Wnt gene family in hematopoiesis: expansion of multilineage progenitor cells. Blood *89*, 3624-3635.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M.*, et al.* (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol *8*, 532-538.

Bachman, K.E., Park, B.H., Rhee, I., Rajagopalan, H., Herman, J.G., Baylin, S.B., Kinzler, K.W., and Vogelstein, B. (2003). Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. Cancer Cell *3*, 89-95.

Bariol, C., Suter, C., Cheong, K., Ku, S.-L., Meagher, A., Hawkins, N., and Ward, R. (2003). The relationship between hypomethylation and CpG island methylation in colorectal neoplasia. Am J Pathol *162*, 1361-1371.

Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F.*, et al.* (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature *445*, 671-675.

Baylin, S.B., and Ohm, J.E. (2006). Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? NatRevCancer *6*, 107-116.

Beier, D., Hau, P., Proescholdt, M., Lohmeier, A., Wischhusen, J., Oefner, P.J., Aigner, L., Brawanski, A., Bogdahn, U., and Beier, C.P. (2007). CD133(+) and CD133(-) glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. Cancer Res *67*, 4010-4015.

Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A., and Weinberg, R.A. (2008). An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat Genet *40*, 499-507.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bhardwaj, G., Murdoch, B., Wu, D., Baker, D.P., Williams, K.P., Chadwick, K., Ling, L.E., Karanu, F.N., and Bhatia, M. (2001). Sonic hedgehog induces the proliferation of primitive human hematopoietic cells via BMP regulation. Nat Immunol *2*, 172-180.

Biniszkiewicz, D., Gribnau, J., Ramsahoye, B., Gaudet, F., Eggan, K., Humpherys, D., Mastrangelo, M.A., Jun, Z., Walter, J., and Jaenisch, R. (2002). Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. MolCell Biol *22*, 2124-2135.

Bird, A., Taggart, M., Frommer, M., Miller, O.J., and Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell *40*, 91-99.

Bjerkvig, R., Tysnes, B.B., Aboody, K.S., Najbauer, J., and Terzis, A.J. (2005). Opinion: the origin of the cancer stem cell: current controversies and new insights. NatRevCancer *5*, 899-904.

Boeke, J., Ammerpohl, O., Kegel, S., Moehren, U., and Renkawitz, R.

Bollati, V., Fabris, S., Pegoraro, V., Ronchetti, D., Mosca, L., Deliliers, G.L., Motta, V., Bertazzi, P.A., Baccarelli, A., and Neri, A. (2009). Differential repetitive DNA methylation in multiple myeloma molecular subgroups. Carcinogenesis *30*, 1330-1335.

Bonnet, D., and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med *3*, 730-737.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K.

Bracken, A.P., Pasini, D., Capra, M., Prosperini, E., Colli, E., and Helin, K. (2003). EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. EMBO J *22*, 5323-5335.

Brothman, A.R., Swanson, G., Maxwell, T.M., Cui, J., Murphy, K.J., Herrick, J., Speights, V.O., Isaac, J., and Rohr, L.R. (2005). Global hypomethylation is common in prostate cancer cells: a quantitative predictor for clinical outcome? Cancer Genet Cytogenet *156*, 31-36.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Chalitchagorn, K., Shuangshoti, S., Hourpai, N., Kongruttanachok, N., Tangkijvanich, P., Thong-ngam, D., Voravud, N., Sriuranpong, V., and Mutirangura, A. (2004). Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. Oncogene *23*, 8841-8846.

Chan, K.Y.K., Ozçelik, H., Cheung, A.N.Y., Ngan, H.Y.S., and Khoo, U.-S. (2002). Epigenetic factors controlling the BRCA1 and BRCA2 genes in sporadic ovarian cancer. Cancer Res *62*, 4151-4156.

Chen, R.Z., Pettersson, U., Beard, C., Jackson-Grusby, L., and Jaenisch, R. (1998). DNA hypomethylation leads to elevated mutation rates. Nature *395*, 89-93.

Chen, T., Hevi, S., Gay, F., Tsujimoto, N., He, T., Zhang, B., Ueda, Y., and Li, E. (2007). Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. NatGenet *39*, 391-396.

Chilukamarri, L., Hancock, A.L., Malik, S., Zabkiewicz, J., Baker, J.A., Greenhough, A., Dallosso, A.R., Huang, T.H.-M., Royer-Pokora, B., Brown, K.W.*, et al.* (2007). Hypomethylation and aberrant expression of the glioma pathogenesis-related 1 gene in Wilms tumors. Neoplasia *9*, 970-978.

Cho, B., Lee, H., Jeong, S., Bang, Y.-J., Lee, H.J., Hwang, K.S., Kim, H.-Y., Lee, Y.-S., Kang, G.H., and Jeoung, D.-I. (2003). Promoter hypomethylation of a novel cancer/testis antigen gene CAGE is correlated with its aberrant expression and is seen in premalignant stage of gastric carcinoma. Biochem Biophys Res Commun *307*, 52-63.

Cho, M., Uemura, H., Kim, S.C., Kawada, Y., Yoshida, K., Hirao, Y., Konishi, N., Saga, S., and Yoshikawa, K. (2001). Hypomethylation of the MN/CA9 promoter and upregulated MN/CA9 expression in human renal cell carcinoma. BrJ Cancer *85*, 563-567.

Choi, I.S., Estecio, M.R., Nagano, Y., Kim, d., White, J.A., Yao, J.C., Issa, J.P., and Rashid, A. (2007). Hypomethylation of LINE-1 and Alu in well-differentiated neuroendocrine tumors (pancreatic endocrine tumors and carcinoid tumors). ModPathol *20*, 802-810.

Choi, S.H., Worswick, S., Byun, H.-M., Shear, T., Soussa, J.C., Wolff, E.M., Douer, D., Garcia-Manero, G., Liang, G., and Yang, A.S. (2009). Changes in DNA methylation of tandem DNA repeats are different from interspersed repeats in cancer. Int J Cancer *125*, 723-729.

Cooper, D.N., and Krawczak, M. (1989). Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. Hum Genet *83*, 181-188.

Costa, F.F., Paixão, V.A., Cavalher, F.P., Ribeiro, K.B., Cunha, I.W., Rinck, J.A., O'Hare, M., Mackay, A., Soares, F.A., Brentani, R.R.*, et al.* (2006). SATR-1 hypomethylation is a common and early event in breast cancer. Cancer Genet Cytogenet *165*, 135-143.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., and Gage, F.H. (2009). L1 retrotransposition in human neural progenitor cells. Nature *460*, 1127-1131.

Counter, C.M., Avilion, A.A., LeFeuvre, C.E., Stewart, N.G., Greider, C.W., Harley, C.B., and Bacchetti, S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. EMBO J *11*, 1921-1929.

Croonquist, P.A., and Van Ness, B. (2005). The polycomb group protein enhancer of zeste homolog 2 (EZH 2) is an oncogene that influences myeloma cell growth and the mutant ras phenotype. Oncogene *24*, 6269-6280.

Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R., and Feinberg, A.P. (2003). Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science *299*, 1753-1755.

Danielian, P.S., White, R., Hoare, S.A., Fawell, S.E., and Parker, M.G. (1993). Identification of residues in the estrogen receptor that confer differential sensitivity to estrogen and hydroxytamoxifen. Mol Endocrinol *7*, 232-240.

Dante, R., nte-Paire, J., Rigal, D., and Roizes, G. (1992). Methylation patterns of long interspersed repeated DNA and alphoid repetitive DNA from human cell lines and tumors. Anticancer Res *12*, 559-563.

Daujat, S., Zeissler, U., Waldmann, T., Happel, N., and Schneider, R. (2005). HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. J Biol Chem *280*, 38090-38095.

Dellino, G.I., Schwartz, Y.B., Farkas, G., McCabe, D., Elgin, S.C.R., and Pirrotta, V. (2004). Polycomb silencing blocks transcription initiation. Mol Cell *13*, 887-893.

Dodge, J.E., Okano, M., Dick, F., Tsujimoto, N., Chen, T., Wang, S., Ueda, Y., Dyson, N., and Li, E. (2005). Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. J Biol Chem *280*, 17986-17991.

Eads, C.A., Danenberg, K.D., Kawakami, K., Saltz, L.B., Danenberg, P.V., and Laird, P.W. (1999). CpG island hypermethylation in human colorectal tumors is not associated with DNA methyltransferase overexpression. Cancer Res *59*, 2302-2306.

Eden, A., Gaudet, F., Waghmare, A., and Jaenisch, R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. Science *300*, 455.

Ehrlich, M. (2002a). DNA hypomethylation, cancer, the immunodeficiency, centromeric region instability, facial anomalies syndrome and chromosomal rearrangements. JNutr *132*, 2424S-2429S.

Ehrlich, M. (2002b). DNA methylation in cancer: too much, but also too little. Oncogene *21*, 5400-5413.

Ehrlich, M. (2006). Cancer-linked DNA hypomethylation and its relationship to hypermethylation. CurrTopMicrobiolImmunol *310*, 251-274.

Ehrlich, M., Buchanan, K.L., Tsien, F., Jiang, G., Sun, B., Uicker, W., Weemaes, C.M., Smeets, D., Sperling, K., Belohradsky, B.H.*, et al.* (2001). DNA methyltransferase 3B mutations linked to the ICF syndrome cause dysregulation of lymphogenesis genes. Hum Mol Genet *10*, 2917-2931.

Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res *10*, 2709-2721.

Ehrlich, M., Jiang, G., Fiala, E., Dome, J.S., Yu, M.C., Long, T.I., Youn, B., Sohn, O.S., Widschwendter, M., Tomlinson, G.E.*, et al.* (2002). Hypomethylation and hypermethylation of DNA in Wilms tumors. Oncogene *21*, 6694-6702.

Ehrlich, M., Woods, C.B., Yu, M.C., Dubeau, L., Yang, F., Campan, M., Weisenberger, D.J., Long, T., Youn, B., Fiala, E.S.*, et al.* (2006). Quantitative analysis of associations between DNA hypermethylation, hypomethylation, and DNMT RNA levels in ovarian tumors. Oncogene *25*, 2636-2645.

Erhardt, S., Su, I.H., Schneider, R., Barton, S., Bannister, A.J., Perez-Burgos, L., Jenuwein, T., Kouzarides, T., Tarakhovsky, A., and Surani, M.A. (2003). Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. Development *130*, 4235-4248.

Esquela-Kerscher, A., and Slack, F.J. (2006). Oncomirs - microRNAs with a role in cancer. NatRevCancer *6*, 259-269.

Esteller, M. (2005). Aberrant DNA methylation as a cancer-inducing mechanism. AnnuRevPharmacolToxicol *45*, 629-656.

Esteller, M. (2006). Epigenetics provides a new generation of oncogenes and tumour-suppressor genes. Br J Cancer *94*, 179-183.

Esteller, M., Hamilton, S.R., Burger, P.C., Baylin, S.B., and Herman, J.G. (1999). Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res *59*, 793-797.

Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell *61*, 759-767.

Feinberg, A.P., Gehrke, C.W., Kuo, K.C., and Ehrlich, M. (1988). Reduced genomic 5-methylcytosine content in human colonic neoplasia. Cancer Res *48*, 1159-1161.

Feinberg, A.P., Ohlsson, R., and Henikoff, S. (2006). The epigenetic progenitor origin of human cancer. NatRevGenet *7*, 21-33.

Feinberg, A.P., and Tycko, B. (2004). The history of cancer epigenetics. NatRevCancer *4*, 143-153.

Feinberg, A.P., and Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature *301*, 89-92.

Feng, B., Ng, J.-H., Heng, J.-C.D., and Ng, H.-H. (2009). Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. Cell Stem Cell *4*, 301-312.

Flanagan, J.M., Funes, J.M., Henderson, S., Wild, L., Carey, N., and Boshoff, C. (2009). Genomics screen in transformed stem cells reveals RNASEH2A, PPAP2C, and ADARB1 as putative anticancer drug targets. Mol Cancer Ther *8*, 249-260.

Florl, A.R., Lower, R., Schmitz-Drager, B.J., and Schulz, W.A. (1999). DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. BrJ Cancer *80*, 1312-1321.

Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K.*, et al.* (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. NatGenet *37*, 391-400.

Fritsch, C., Brown, J.L., Kassis, J.A., and Müller, J. (1999). The DNA-binding polycomb group protein pleiohomeotic mediates silencing of a Drosophila homeotic gene. Development *126*, 3905-3913.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. ProcNatl AcadSciUSA *89*, 1827-1831.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L., and Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nat Genet *24*, 88-91.

Funes, J.M., Quintero, M., Henderson, S., Martinez, D., Qureshi, U., Westwood, C., Clements, M.O., Bourboulia, D., Pedley, R.B., Moncada, S.*, et al.* (2007). Transformation of human mesenchymal stem cells increases their dependency on oxidative phosphorylation for energy production. ProcNatl AcadSciUSA *104*, 6223-6228.

Galli, R., Binda, E., Orfanelli, U., Cipelletti, B., Gritti, A., De Vitis, S., Fiocco, R., Foroni, C., Dimeco, F., and Vescovi, A. (2004). Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. Cancer Res *64*, 7011-7021.

Gama-Sosa, M.A., Midgett, R.M., Slagel, V.A., Githens, S., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983a). Tissue-specific differences in DNA methylation in various mammals. BiochimBiophysActa *740*, 212-219.

Gama-Sosa, M.A., Slagel, V.A., Trewyn, R.W., Oxenhandler, R., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983b). The 5-methylcytosine content of DNA from human tumors. Nucleic Acids Res *11*, 6883-6894.

Gama-Sosa, M.A., Wang, R.Y., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983c). The 5-methylcytosine content of highly repeated sequences in human DNA. Nucleic Acids Res *11*, 3087-3095.

Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., and Jaenisch, R. (2003). Induction of tumors in mice by genomic hypomethylation. Science *300*, 489-492.

Gazin, C., Wajapeyee, N., Gobeil, S., Virbasius, C.M., and Green, M.R. (2007). An elaborate pathway required for Ras-mediated epigenetic silencing. Nature *449*, 1073-1077.

Ginestier, C., Hur, M.H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., Jacquemier, J., Viens, P., Kleer, C.G., Liu, S.*, et al.* (2007). ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell *1*, 555-567.

Gjerset, R.A., and Martin, D.W. (1982). Presence of a DNA demethylating activity in the nucleus of murine erythroleukemic cells. J Biol Chem *257*, 8581-8583.

Goelz, S.E., Vogelstein, B., Hamilton, S.R., and Feinberg, A.P. (1985). Hypomethylation of DNA from benign and malignant human colon neoplasms. Science *228*, 187-190.

Gregorieff, A., and Clevers, H. (2005). Wnt signaling in the intestinal epithelium: from endoderm to cancer. Genes Dev *19*, 877-890.

Gupta, A., Godwin, A.K., Vanderveer, L., Lu, A., and Liu, J. (2003). Hypomethylation of the synuclein gamma gene CpG island promotes its aberrant expression in breast carcinoma and ovarian carcinoma. Cancer Res *63*, 664-673.

Gupta, P.B., Proia, D., Cingoz, O., Weremowicz, J., Naber, S.P., Weinberg, R.A., and Kuperwasser, C. (2007). Systemic stromal effects of estrogen promote the growth of estrogen receptor-negative cancers. Cancer Res *67*, 2062-2071.

Habib, M., Fares, F., Bourgeois, C.A., Bella, C., Bernardino, J., Hernandez-Blazquez, F., de, C., and Niveleau, A. (1999). DNA global hypomethylation in EBV-transformed interphase nuclei. ExpCell Res *249*, 46-53.

Haldar, M., Hancock, J.D., Coffin, C.M., Lessnick, S.L., and Capecchi, M.R. (2007). A conditional mouse model of synovial sarcoma: insights into a myogenic origin. Cancer Cell *11*, 375-388.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57-70.

Havas, K., Flaus, A., Phelan, M., Kingston, R., Wade, P.A., Lilley, D.M., and Owen-Hughes, T. (2000). Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. Cell *103*, 1133-1142.

Hayflick, L. (1965). The Limited In Vitro Lifetime Of Human Diploid Cell Strains. Exp Cell Res *37*, 614-636.

Hayflick, L., and Moorhead, P.S. (1961). The serial cultivation of human diploid cell strains. Exp Cell Res *25*, 585-621.

He, X.C., Zhang, J., and Li, L. (2005). Cellular and molecular regulation of hematopoietic and intestinal stem cell behavior. Ann N Y Acad Sci *1049*, 28-38.

Helman, L.J., and Meltzer, P. (2003). Mechanisms of sarcoma development. Nat Rev Cancer *3*, 685-694.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V.A., and Bird, A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev *15*, 710-723.

Herman, J.G., and Baylin, S.B. (2003). Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med *349*, 2042-2054.

Herman, J.G., Umar, A., Polyak, K., Graff, J.R., Ahuja, N., Issa, J.P., Markowitz, S., Willson, J.K., Hamilton, S.R., Kinzler, K.W.*, et al.* (1998). Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci USA *95*, 6870-6875.

Hernandez, R., Frady, A., Zhang, X.Y., Varela, M., and Ehrlich, M. (1997). Preferential induction of chromosome 1 multibranched figures and whole-arm

deletions in a human pro-B cell line treated with 5-azacytidine or 5-azadeoxycytidine. Cytogenet Cell Genet *76*, 196-201.

Hernandez-Blazquez, F.J., Habib, M., Dumollard, J.M., Barthelemy, C., Benchaib, M., de Capoa, A., and Niveleau, A. (2000). Evaluation of global DNA hypomethylation in human colon cancer tissues by immunohistochemistry and image analysis. Gut *47*, 689-693.

Hochedlinger, K., Blelloch, R., Brennan, C., Yamada, Y., Kim, M., Chin, L., and Jaenisch, R. (2004). Reprogramming of a melanoma genome by nuclear transplantation. Genes Dev *18*, 1875-1885.

Houghton, J., Stoicov, C., Nomura, S., Rogers, A.B., Carlson, J., Li, H., Cai, X., Fox, J.G., Goldenring, J.R., and Wang, T.C. (2004). Gastric cancer originating from bone marrow-derived cells. Science *306*, 1568-1571.

Hsieh, C.L., and Lieber, M.R. (1992). CpG methylated minichromosomes become inaccessible for V(D)J recombination after undergoing replication. EMBO J *11*, 315-325.

Huntly, B.J.P., Shigematsu, H., Deguchi, K., Lee, B.H., Mizuno, S., Duclos, N., Rowan, R., Amaral, S., Curley, D., Williams, I.R.*, et al.* (2004). MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. Cancer Cell *6*, 587-596.

Iacopetta, B., Grieu, F., Phillips, M., Ruszkiewicz, A., Moore, J., Minamoto, T., and Kawakami, K. (2007). Methylation levels of LINE-1 repeats and CpG island loci are inversely related in normal colonic mucosa. Cancer Sci *98*, 1454-1460.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E.*, et al.* (2001). Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat Genet *27*, 31-39.

Jacobs, J.J., Kieboom, K., Marino, S., DePinho, R.A., and van Lohuizen, M. (1999a). The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. Nature *397*, 164-168.

Jacobs, J.J., Scheijen, B., Voncken, J.W., Kieboom, K., Berns, A., and van Lohuizen, M. (1999b). Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. Genes Dev *13*, 2678-2690.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet *33 Suppl*, 245-254.

Jamieson, C.H.M., Ailles, L.E., Dylla, S.J., Muijtjens, M., Jones, C., Zehnder, J.L., Gotlib, J., Li, K., Manz, M.G., Keating, A.*, et al.* (2004). Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. N Engl J Med *351*, 657-667.

Jang, S.J., Soria, J.C., Wang, L., Hassan, K.A., Morice, R.C., Walsh, G.L., Hong, W.K., and Mao, L. (2001). Activation of melanoma antigen tumor antigens occurs early in lung carcinogenesis. Cancer Res *61*, 7959-7963.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Ji, W., Hernandez, R., Zhang, X.Y., Qu, G.Z., Frady, A., Varela, M., and Ehrlich, M. (1997). DNA demethylation and pericentromeric rearrangements of chromosome 1. Mutat Res *379*, 33-41.

Jin, S.-G., Guo, C., and Pfeifer, G.P. (2008). GADD45A does not promote DNA demethylation. PLoS Genet *4*, e1000013.

Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. NatRevGenet *3*, 415-428.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet *19*, 187-191.

Jurgens, B., Schmitz-Drager, B.J., and Schulz, W.A. (1996). Hypomethylation of L1 LINE sequences prevailing in human urothelial carcinoma. Cancer Res *56*, 5698-5703.

Kaludov, N.K., and Wolffe, A.P. (2000). MeCP2 driven transcriptional repression in vitro: selectivity for methylated DNA, action at a distance and contacts with the basal transcription machinery. Nucleic Acids Res *28*, 1921-1928.

Kanai, Y., Ushijima, S., Kondo, Y., Nakanishi, Y., and Hirohashi, S. (2001). DNA methyltransferase expression and DNA methylation of CPG islands and peri-centromeric satellite regions in human colorectal and stomach cancers. Int J Cancer *91*, 205-212.

Kane, M.F., Loda, M., Gaida, G.M., Lipman, J., Mishra, R., Goldman, H., Jessup, J.M., and Kolodner, R. (1997). Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res *57*, 808-811.

Kaneda, A., Tsukamoto, T., Takamura-Enya, T., Watanabe, N., Kaminishi, M., Sugimura, T., Tatematsu, M., and Ushijima, T. (2004). Frequent hypomethylation in multiple promoter CpG islands is associated with global hypomethylation, but not with frequent promoter hypermethylation. Cancer Sci *95*, 58-64.

Karanu, F.N., Murdoch, B., Gallacher, L., Wu, D.M., Koremoto, M., Sakano, S., and Bhatia, M. (2000). The notch ligand jagged-1 represents a novel growth factor of human hematopoietic stem cells. J Exp Med *192*, 1365-1372.

Karnoub, A.E., Dash, A.B., Vo, A.P., Sullivan, A., Brooks, M.W., Bell, G.W., Richardson, A.L., Polyak, K., Tubo, R., and Weinberg, R.A. (2007). Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. Nature *449*, 557-563.

Kazazian, H.H., and Moran, J.V. (1998). The impact of L1 retrotransposons on the human genome. Nat Genet *19*, 19-24.

Kelly, L.M., and Gilliland, D.G. (2002). Genetics of myeloid leukemias. Annual review of genomics and human genetics *3*, 179-198.

Kelly, P.N., Dakic, A., Adams, J.M., Nutt, S.L., and Strasser, A. (2007). Tumor growth need not be driven by rare cancer stem cells. Science *317*, 337.

Kempski, H.M., Austin, N., Chatters, S.J., Toomey, S.M., Chalker, J., Anderson, J., and Sebire, N.J. (2006). Previously unidentified complex cytogenetic changes found in a pediatric case of solid-pseudopapillary neoplasm of the pancreas. Cancer Genet Cytogenet *164*, 54-60.

Kim, C.F.B., Jackson, E.L., Woolfenden, A.E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R.T., and Jacks, T. (2005). Identification of bronchioalveolar stem cells in normal lung and lung cancer. Cell *121*, 823-835.

Kim, Y.-I. (2005). Nutritional epigenetics: impact of folate deficiency on DNA methylation and colon cancer susceptibility. J Nutr *135*, 2703-2709.

Kiyono, T., Foster, S.A., Koop, J.I., McDougall, J.K., Galloway, D.A., and Klingelhutz, A.J. (1998). Both Rb/p16INK4a inactivation and telomerase activity are required to immortalize human epithelial cells. Nature *396*, 84-88.

Kleer, C.G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S.A., Ghosh, D., Sewalt, R.G., Otte, A.P., Hayes, D.F.*, et al.* (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. ProcNatl AcadSciUSA *100*, 11606-11611.

Knudson, A.G., Strong, L.C., and Anderson, D.E. (1973). Heredity and cancer in man. Progress in medical genetics *9*, 113-158.

Kochanek, S., Renz, D., and Doerfler, W. (1995). Transcriptional silencing of human Alu sequences and inhibition of protein binding in the box B regulatory elements by 5'-CG-3' methylation. FEBS Lett *360*, 115-120.

Kondo, Y., Shen, L., Cheng, A.S., Ahmed, S., Boumber, Y., Charo, C., Yamochi, T., Urano, T., Furukawa, K., Kwabi-Addo, B.*, et al.* (2008). Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. Nat Genet *40*, 741-750.

Kondo, Y., Shen, L., Yan, P.S., Huang, T.H.-M., and Issa, J.-P.J. (2004). Chromatin immunoprecipitation microarrays for identification of genes silenced by histone H3 lysine 9 methylation. Proc Natl Acad Sci USA *101*, 7398-7403.

Kopelovich, L., Crowell, J.A., and Fay, J.R. (2003). The epigenome as a target for cancer chemoprevention. JNatlCancer Inst *95*, 1747-1757.

Krivtsov, A.V., Twomey, D., Feng, Z., Stubbs, M.C., Wang, Y., Faber, J., Levine, J.E., Wang, J., Hahn, W.C., Gilliland, D.G.*, et al.* (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature *442*, 818-822.

Kuo, K.C., McCune, R.A., Gehrke, C.W., Midgett, R., and Ehrlich, M. (1980). Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. Nucleic Acids Res *8*, 4763-4776.

Kuperwasser, C., Chavarria, T., Wu, M., Magrane, G., Gray, J.W., Carey, L., Richardson, A., and Weinberg, R.A. (2004). Reconstruction of functionally normal and malignant human breast tissues in mice. Proc Natl Acad Sci USA *101*, 4966-4971.

Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol Cell *14*, 183-193.

Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T.S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P.*, et al.* (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. Proc Natl Acad Sci USA *102*, 1859-1864.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev *16*, 2893-2905.

Laird, P.W., and Jaenisch, R. (1996). The role of DNA methylation in cancer genetic and epigenetics. AnnuRev Genet *30*, 441-464.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A., and Dick, J.E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature *367*, 645-648.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.-i., *et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell *71*, 865-873.

Lessard, J., and Sauvageau, G. (2003). Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. Nature *423*, 255-260.

Levine, S.S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., and Kingston, R.E. (2002). The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. Mol Cell Biol *22*, 6070-6078.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li, S., Chiang, T.-c., Richard-Davis, G., Barrett, J.C., and Mclachlan, J.A. (2003). DNA hypomethylation and imbalanced expression of DNA methyltransferases (DNMT1, 3A, and 3B) in human uterine leiomyoma. Gynecol Oncol *90*, 123-130.

Lin, C.H., Hsieh, S.Y., Sheen, I.S., Lee, W.C., Chen, T.C., Shyu, W.C., and Liaw, Y.F. (2001). Genome-wide hypomethylation in hepatocellular carcinogenesis. Cancer Res *61*, 4238-4243.

Liu, H., Liu, W., Wu, Y., Zhou, Y., Xue, R., Luo, C., Wang, L., Zhao, W., Jiang, J.-D., and Liu, J. (2005). Loss of epigenetic control of synuclein-gamma gene as a molecular indicator of metastasis in a wide range of human cancers. Cancer Res *65*, 7635-7643.

Lotem, J., and Sachs, L. (2002). Epigenetics wins over genetics: induction of differentiation in tumor cells. Semin Cancer Biol *12*, 339-346.

Lund, P., Weisshaupt, K., Mikeska, T., Jammas, D., Chen, X., Kuban, R.J., Ungethum, U., Krapfenbauer, U., Herzel, H.P., Schafer, R., *et al.* (2006). Oncogenic HRAS suppresses clusterin expression through promoter hypermethylation. Oncogene *25*, 4890-4903.

MacLeod, A.R., Rouleau, J., and Szyf, M. (1995). Regulation of DNA methylation by the Ras signaling pathway. J Biol Chem *270*, 11327-11337.

Malik, K., Salpekar, A., Hancock, A., Moorwood, K., Jackson, S., Charles, A., and Brown, K.W. (2000). Identification of differential methylation of the WT1 antisense regulatory region and relaxation of imprinting in Wilms' tumor. Cancer Res *60*, 2356-2360.

Maloisel, L., and Rossignol, J.L. (1998). Suppression of crossing-over by DNA methylation in Ascobolus. Genes Dev *12*, 1381-1389.

Marmorstein, R., and Roth, S.Y. (2001). Histone acetyltransferases: function, structure, and catalysis. Curr Opin Genet Dev *11*, 155-161.

Martin, C., Cao, R., and Zhang, Y. (2006). Substrate preferences of the EZH2 histone methyltransferase complex. J Biol Chem *281*, 8365-8370.

Martin, V., Ribieras, S., Song-Wang, X.G., Lasne, Y., Frappart, L., Rio, M.C., and Dante, R.

Mattioni, T., Louvion, J.F., and Picard, D. (1994). Regulation of protein activities by fusion to steroid binding domains. Methods Cell Biol *43 Pt A*, 335-352.

McGarvey, K.M., Greene, E., Fahrner, J.A., Jenuwein, T., and Baylin, S.B. (2007). DNA Methylation and Complete Transcriptional Silencing of Cancer Genes Persist after Depletion of EZH2. Cancer Res *67*, 5097-5102.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B.*, et al.* (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766-770.

Mihaly, J., Mishra, R.K., and Karch, F. (1998). A conserved sequence motif in Polycomb-response elements. Mol Cell *1*, 1065-1066.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res *16*, 1215.

Min, J., Zhang, Y., and Xu, R.-M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev *17*, 1823-1828.

Montagna, M., Santacatterina, M., Torri, A., Menin, C., Zullato, D., Chieco-Bianchi, L., and D'Andrea, E. (1999). Identification of a 3 kb Alu-mediated BRCA1 gene rearrangement in two breast/ovarian cancer families. Oncogene *18*, 4160-4165.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J Biol Chem *279*, 52353-52360.

Morgan, H.D., Santos, F., Green, K., Dean, W., and Reik, W. (2005). Epigenetic reprogramming in mammals. HumMol Genet *14 Spec No 1*, R47-R58.

Morin, P.J., Sparks, A.B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B., and Kinzler, K.W. (1997). Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. Science *275*, 1787-1790.

Moser, A.R., Dove, W.F., Roth, K.A., and Gordon, J.I. (1992). The Min (multiple intestinal neoplasia) mutation: its effect on gut epithelial cell differentiation and interaction with a modifier system. J Cell Biol *116*, 1517-1526.

Müller, J., and Kassis, J.A. (2006). Polycomb response elements and targeting of Polycomb group proteins in Drosophila. Curr Opin Genet Dev *16*, 476-484.

Nagai, H., Kim, Y.S., Yasuda, T., Ohmachi, Y., Yokouchi, H., Monden, M., Emi, M., Konishi, N., Nogami, M., Okumura, K.*, et al.* (1999). A novel sperm-specific hypomethylation sequence is a demethylation hotspot in human hepatocellular carcinomas. Gene *237*, 15-20.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2007). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Nan, X., Tate, P., Li, E., and Bird, A. (1996). DNA methylation specifies chromosomal localization of MeCP2. Mol Cell Biol *16*, 414-421.

Narayan, A., Ji, W., Zhang, X.Y., Marrogi, A., Graff, J.R., Baylin, S.B., and Ehrlich, M. (1998). Hypomethylation of pericentromeric DNA in breast adenocarcinomas. IntJ Cancer *77*, 833-838.

Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., and Bird, A. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. Nat Genet *23*, 58-61.

O'Brien, C.A., Pollett, A., Gallinger, S., and Dick, J.E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature *445*, 106-110.

O'Neill, R.J., O'Neill, M.J., and Graves, J.A. (1998). Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. Nature *393*, 68-72.

Ogasawara, S., Maesawa, C., Yamamoto, M., Akiyama, Y., Wada, K., Fujisawa, K., Higuchi, T., Tomisawa, Y., Sato, N., Endo, S.*, et al.* (2004). Disruption of cell-type-specific methylation at the Maspin gene promoter is frequently involved in undifferentiated thyroid cancers. Oncogene *23*, 1117-1124.

Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L., Schuebel, K.E., Cope, L., Mohammad, H.P., Chen, W., Daniel, V.C., Yu, W.*, et al.* (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet *39*, 237-242.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Okano, M., Xie, S., and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet *19*, 219-220.

Ooi, S.K., and Bestor, T.H. (2008). The colorful history of active DNA demethylation. Cell *133*, 1145-1148.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr Biol *10*, 475-478.

Pardal, R., Clarke, M.F., and Morrison, S.J. (2003). Applying the principles of stem-cell biology to cancer. NatRevCancer *3*, 895-902.

Park, I.-k., Qian, D., Kiel, M., Becker, M.W., Pihalja, M., Weissman, I.L., Morrison, S.J., and Clarke, M.F. (2003). Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. Nature *423*, 302-305.

Pattamadilok, J., Huapai, N., Rattanatanyong, P., Vasurattana, A., Triratanachat, S., Tresukosol, D., and Mutirangura, A. (2008). LINE-1 hypomethylation level as a potential prognostic factor for epithelial ovarian cancer. Int J Gynecol Cancer *18*, 711-717.

Paz, M.F., Avila, S., Fraga, M.F., Pollan, M., Capella, G., Peinado, M.A., Sanchez-Cespedes, M., Herman, J.G., and Esteller, M. (2002). Germ-line variants in methyl-group metabolism genes and susceptibility to DNA methylation in normal tissues and human primary tumors. Cancer Res *62*, 4519-4524.

Pereira, D.S., Dorrell, C., Ito, C.Y., Gan, O.I., Murdoch, B., Rao, V.N., Zou, J.P., Reddy, E.S., and Dick, J.E. (1998). Retroviral transduction of TLS-ERG initiates a leukemogenic program in normal human hematopoietic cells. Proc Natl Acad Sci USA *95*, 8239-8244.

Polyak, K., and Hahn, W.C. (2006). Roots and stems: stem cells in cancer. NatMed *12*, 296-300.

Powell, S.M., Zilz, N., Beazer-Barclay, Y., Bryan, T.M., Hamilton, S.R., Thibodeau, S.N., Vogelstein, B., and Kinzler, K.W. (1992). APC mutations occur early during colorectal tumorigenesis. Nature *359*, 235-237.

Pradhan, S., Bacolla, A., Wells, R.D., and Roberts, R.J. (1999). Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. J Biol Chem *274*, 33002-33010.

Prak, E.T., and Kazazian, H.H. (2000). Mobile elements and the human genome. Nat Rev Genet *1*, 134-144.

Qiu, J. (2006). Epigenetics: unfinished symphony. Nature *441*, 143-145.

Qu, G., Dubeau, L., Narayan, A., Yu, M.C., and Ehrlich, M. (1999). Satellite DNA hypomethylation vs. overall genomic hypomethylation in ovarian epithelial tumors of different malignant potential. MutatRes *423*, 91-101.

Quintana, E., Shackleton, M., Sabel, M.S., Fullen, D.R., Johnson, T.M., and Morrison, S.J. (2008). Efficient tumour formation by single human melanoma cells. Nature *456*, 593-598.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell *135*, 1201-1212.

Rangarajan, A., Hong, S.J., Gifford, A., and Weinberg, R.A. (2004). Species- and cell type-specific requirements for cellular transformation. Cancer Cell *6*, 171-183.

Rauch, T.A., Zhong, X., Wu, X., Wang, M., Kernstine, K.H., Wang, Z., Riggs, A.D., and Pfeifer, G.P. (2008). High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci USA *105*, 252-257.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089-1093.

Reya, T., and Clevers, H. (2005). Wnt signalling in stem cells and cancer. Nature *434*, 843-850.

Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. Nature *414*, 105-111.

Rhee, I., Jair, K.W., Yen, R.W., Lengauer, C., Herman, J.G., Kinzler, K.W., Vogelstein, B., Baylin, S.B., and Schuebel, K.E. (2000). CpG methylation is maintained in human cancer cells lacking DNMT1. Nature *404*, 1003-1007.

Ribieras, S., Song-Wang, X.G., Martin, V., Lointier, P., Frappart, L., and Dante, R. (1994). Human breast and colon cancers exhibit alterations of DNA methylation patterns at several DNA segments on chromosomes 11p and 17p. J Cell Biochem *56*, 86-96.

Richards, K.L., Zhang, B., Baggerly, K.A., Colella, S., Lang, J.C., Schuller, D.E., and Krahe, R. (2009). Genome-wide hypomethylation in head and neck cancer is more pronounced in HPV-negative tumors and is associated with genomic instability. PLoS ONE *4*, e4941.

Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet *38*, 413-443.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., *et al*. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L., and Wolffe, A.P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nat Genet *25*, 338-342.

Robertson, K.D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F.A., and Jones, P.A. (1999). The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. Nucleic Acids Res *27*, 2291-2298.

Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Castillejo, J.A., Navarro, G., San Jose-Eneriz, E., Garate, L., Cordeu, L., Cervantes, F., Prosper, F., *et al.* (2008). Repetitive DNA hypomethylation in the advanced phase of chronic myeloid leukemia. LeukRes *32*, 487-490.

Romanov, S.R., Kozakiewicz, B.K., Holst, C.R., Stampfer, M.R., Haupt, L.M., and Tlsty, T.D. (2001). Normal human mammary epithelial cells spontaneously escape senescence and acquire genomic changes. Nature *409*, 633-637.

Rouleau, J., MacLeod, A.R., and Szyf, M. (1995). Regulation of the DNA methyltransferase by the Ras-AP-1 signaling pathway. J Biol Chem *270*, 1595-1601.

Rountree, M.R., Bachman, K.E., and Baylin, S.B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat Genet *25*, 269-277.

Rubin, C.M., VandeVoort, C.A., Teplitz, R.L., and Schmid, C.W. (1994). Alu repeated DNAs are differentially methylated in primate germ cells. Nucleic Acids Res *22*, 5121-5127.

Sachs, L. (1986). Hematopoietic growth and differentiation factors and the reversibility of malignancy: cell differentiation and by-passing of genetic defects in leukemia. Medical oncology and tumor pharmacotherapy *3*, 165-176.

Saito, Y., Kanai, Y., Sakamoto, M., Saito, H., Ishii, H., and Hirohashi, S. (2001). Expression of mRNA for DNA methyltransferases and methyl-CpG-binding proteins and DNA methylation status on CpG islands and pericentromeric satellite regions during human hepatocarcinogenesis. Hepatology *33*, 561-568.

Sakatani, T., Kaneda, A., Iacobuzio-Donahue, C.A., Carter, M.G., de Boom Witzel, S., Okano, H., Ko, M.S.H., Ohlsson, R., Longo, D.L., and Feinberg, A.P. (2005). Loss of imprinting of Igf2 alters intestinal maturation and tumorigenesis in mice. Science *307*, 1976-1978.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. DevBiol *241*, 172-182.

Santourlidis, S., Florl, A., Ackermann, R., Wirtz, H.C., and Schulz, W.A. (1999). High frequency of alterations in DNA methylation in adenocarcinoma of the prostate. Prostate *39*, 166-174.

Sato, N., Maitra, A., Fukushima, N., van Heek, N.T., Matsubayashi, H., Iacobuzio-Donahue, C.A., Rosty, C., and Goggins, M. (2003). Frequent hypomethylation of multiple genes overexpressed in pancreatic ductal adenocarcinoma. Cancer Res *63*, 4158-4166.

Schatton, T., Murphy, G.F., Frank, N.Y., Yamaura, K., Waaga-Gasser, A.M., Gasser, M., Zhan, Q., Jordan, S., Duncan, L.M., Weishaupt, C*., et al.* (2008). Identification of cells initiating human melanomas. Nature *451*, 345-349.

Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B.E*., et al.* (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet *39*, 232-236.

Schulz, W.A., Elo, J.P., Florl, A.R., Pennanen, S., Santourlidis, S., Engers, R., Buchardt, M., Seifert, H.-H., and Visakorpi, T.

Schwienbacher, C., Gramantieri, L., Scelfo, R., Veronese, A., Calin, G.A., Bolondi, L., Croce, C.M., Barbanti-Brodano, G., and Negrini, M. (2000). Gain of imprinting at chromosome 11p15: A pathogenetic mechanism identified in human hepatocarcinomas. Proc Natl Acad Sci USA *97*, 5445-5449.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C.T., Bender, W., and Kingston, R.E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell *98*, 37-46.

Shapiro, R., Braverman, B., Louis, J.B., and Servis, R.E. (1973). Nucleic acid reactivity and conformation. II. Reaction of cytosine and uracil with sodium bisulfite. J Biol Chem *248*, 4060-4064.

Sharrard, R.M., Royds, J.A., Rogers, S., and Shorthouse, A.J. (1992). Patterns of methylation of the c-myc gene in human colorectal cancer progression. BrJ Cancer *65*, 667-672.

Shen, L., Fang, J., Qiu, D., Zhang, T., Yang, J., Chen, S., and Xiao, S. (1998). Correlation between DNA methylation and pathological changes in human hepatocellular carcinoma. Hepatogastroenterology *45*, 1753-1759.

Simon, J.A., and Lange, C.A. (2008). Roles of the EZH2 histone methyltransferase in cancer epigenetics. MutatRes *647*, 21-29.

Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol *10*, 697-708.

Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. Nature *432*, 396-401.

Soares, J., Pinto, A.E., Cunha, C.V., André, S., Barão, I., Sousa, J.M., and Cravo, M. (1999). Global DNA hypomethylation in breast carcinoma: correlation with prognostic factors and tumor progression. Cancer *85*, 112-118.

Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. NatRevCancer *6*, 846-856.

Squazzo, S.L., O'Geen, H., Komashko, V.M., Krig, S.R., Jin, V.X., Jang, S.-w., Margueron, R., Reinberg, D., Green, R., and Farnham, P.J. (2006). Suz12 binds to silenced regions of the genome in a cell-type-specific manner. Genome Res *16*, 890-900.

Sterner, D.E., and Berger, S.L. (2000). Acetylation of histones and transcription-related factors. Microbiol Mol Biol Rev *64*, 435-459.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA *102*, 15545-15550.

Sullivan, M.J., Taniguchi, T., Jhee, A., Kerr, N., and Reeve, A.E. (1999). Relaxation of IGF2 imprinting in Wilms tumours associated with specific changes in IGF2 methylation. Oncogene *18*, 7527-7534.

Suter, C.M., Martin, D.I., and Ward, R.L. (2004). Hypomethylation of L1 retrotransposons in colorectal cancer and adjacent normal tissue. International journal of colorectal disease *19*, 95-101.

Suzuki, H., Watkins, D.N., Jair, K.-W., Schuebel, K.E., Markowitz, S.D., Chen, W.D., Pretlow, T.P., Yang, B., Akiyama, Y., Van Engeland, M.*, et al.* (2004). Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. Nat Genet *36*, 417-422.

Szyf, M., Pakneshan, P., and Rabbani, S.A. (2004). DNA demethylation and cancer: therapeutic implications. Cancer Lett *211*, 133-143.

Szyf, M., Theberge, J., and Bozovic, V. (1995). Ras induces a general DNA demethylation activity in mouse embryonal P19 cells. J Biol Chem *270*, 12690-12696.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takai, D., Gonzales, F.A., Tsai, Y.C., Thayer, M.J., and Jones, P.A. (2001). Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. Hum Mol Genet *10*, 2619-2626.

Takai, D., Yagi, Y., Habib, N., Sugimura, T., and Ushijima, T. (2000). Hypomethylation of LINE1 retrotransposon in human hepatocellular carcinomas, but not in surrounding liver cirrhosis. JpnJ Clin Oncol *30*, 306-309.

Taketo, M.M. (2004). Shutting down Wnt signal-activated cancer. Nat Genet *36*, 320-322.

Tao, L., Yang, S., Xie, M., Kramer, P.M., and Pereira, M.A. (2000). Hypomethylation and overexpression of c-jun and c-myc protooncogenes and increased DNA methyltransferase activity in dichloroacetic and trichloroacetic acid-promoted mouse liver tumors. Cancer Lett *158*, 185-193.

Tazi, J., and Bird, A. (1990). Alternative chromatin structure at CpG islands. Cell *60*, 909-920.

Thiagalingam, S., Cheng, K.-H., Lee, H.J., Mineva, N., Thiagalingam, A., and Ponte, J.F. (2003). Histone deacetylases: unique players in shaping the epigenetic histone code. Ann N Y Acad Sci *983*, 84-100.

Trinh, B.N., Long, T.I., and Laird, P.W. (2001). DNA methylation analysis by MethyLight technology. Methods *25*, 456-462.

Tsuda, H., Takarabe, T., Kanai, Y., Fukutomi, T., and Hirohashi, S. (2002). Correlation of DNA hypomethylation at pericentromeric heterochromatin regions of chromosomes 16 and 1 with histological features and chromosomal abnormalities of human breast carcinomas. Am J Pathol *161*, 859-866.

Tuck-Muller, C.M., Narayan, A., Tsien, F., Smeets, D.F., Sawyer, J., Fiala, E.S., Sohn, O.S., and Ehrlich, M. (2000). DNA hypomethylation and unusual chromosome instability in cell lines from ICF syndrome patients. Cytogenet Cell Genet *89*, 121-128.

Turner, B.M. (2002). Cellular memory and the histone code. Cell *111*, 285-291.

van de Wetering, M., Sancho, E., Verweij, C., de Lau, W., Oving, I., Hurlstone, A., van der Horn, K., Batlle, E., Coudreuse, D., Haramis, A.P.*, et al.* (2002). The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell *111*, 241-250.

Van Den Berg, D.J., Sharma, A.K., Bruno, E., and Hoffman, R. (1998). Role of members of the Wnt gene family in human hematopoiesis. Blood *92*, 3189-3202.

van Staveren, W.C., Solis, D.Y., Hebrant, A., Detours, V., Dumont, J.E., and Maenhaut, C. (2009). Human cancer cell lines: Experimental models for cancer cells in situ? For cancer stem cells? BiochimBiophysActa *1795*, 92-103.

Varambally, S., Cao, Q., Mani, R.S., Shankar, S., Wang, X., Ateeq, B., Laxman, B., Cao, X., Jing, X., Ramnarayanan, K.*, et al.* (2008). Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science *322*, 1695-1699.

Varambally, S., Dhanasekaran, S.M., Zhou, M., Barrette, T.R., Kumar-Sinha, C., Sanda, M.G., Ghosh, D., Pienta, K.J., Sewalt, R.G.A.B., Otte, A.P.*, et al.* (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature *419*, 624-629.

Varnum-Finney, B., Xu, L., Brashem-Stein, C., Nourigat, C., Flowers, D., Bakkour, S., Pear, W.S., and Bernstein, I.D. (2000). Pluripotent, cytokine-dependent, hematopoietic stem cells are immortalized by constitutive Notch1 signaling. Nat Med *6*, 1278-1281.

Vertino, P.M., Yen, R.W., Gao, J., and Baylin, S.B. (1996). De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5-)-methyltransferase. Mol Cell Biol *16*, 4555-4565.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van, E., Bernard, D., Vanderwinden, J.M.*, et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871-874.

Visvader, J.E., and Lindeman, G.J. (2008). Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. Nat Rev Cancer *8*, 755-768.

Wade, P.A., Gegonne, A., Jones, P.L., Ballestar, E., Aubry, F., and Wolffe, A.P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nat Genet *23*, 62-66.

Wagner, J., Claverie, N., and Danzin, C. (1984). A rapid high-performance liquid chromatographic procedure for the simultaneous determination of methionine, ethionine, S-adenosylmethionine, S-adenosylethionine, and the natural polyamines in rat tissues. Anal Biochem *140*, 108-116.

Walsh, C.P., Chaillet, J.R., and Bestor, T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet *20*, 116-117.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature *431*, 873-878.

Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. NatGenet *37*, 853-862.

Weinmann, A.S., Bartley, S.M., Zhang, T., Zhang, M.Q., and Farnham, P.J. (2001). Use of chromatin immunoprecipitation to clone novel E2F target promoters. Mol Cell Biol *21*, 6820-6832.

Weinstein, I.B. (2002). Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science *297*, 63-64.

Weisenberger, D.J., Campan, M., Long, T.I., Kim, M., Woods, C., Fiala, E., Ehrlich, M., and Laird, P.W. (2005). Analysis of repetitive element DNA methylation by MethyLight. Nucleic Acids Res *33*, 6823-6836.

Weiss, A., Keshet, I., Razin, A., and Cedar, H. (1996). DNA demethylation in vitro: involvement of RNA. Cell *86*, 709-718.

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D.J., Campan, M., Young, J., Jacobs, I.*, et al.* (2007). Epigenetic stem cell signature in cancer. Nat Genet *39*, 157-158.

Wilson, A.S., Power, B.E., and Molloy, P.L. (2007). DNA hypomethylation and human diseases. BiochimBiophysActa *1775*, 138-162.

Wong, N., Lam, W.C., Lai, P.B., Pang, E., Lau, W.Y., and Johnson, P.J. (2001). Hypomethylation of chromosome 1 heterochromatin DNA correlates with q-arm copy gain in human hepatocellular carcinoma. Am J Pathol *159*, 465-471.

Woodson, K., Flood, A., Green, L., Tangrea, J.A., Hanson, J., Cash, B., Schatzkin, A., and Schoenfeld, P. (2004). Loss of insulin-like growth factor-II imprinting and the presence of screen-detected colorectal adenomas in women. J Natl Cancer Inst *96*, 407-410.

Wright, M.H., Calcagno, A.M., Salcido, C.D., Carlson, M.D., Ambudkar, S.V., and Varticovski, L. (2008). Brca1 breast tumors contain distinct CD44+/CD24- and CD133+ cells with cancer stem cell characteristics. Breast Cancer Res *10*, R10.

Yamada, Y., Jackson-Grusby, L., Linhart, H., Meissner, A., Eden, A., Lin, H., and Jaenisch, R. (2005). Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis. Proc Natl Acad Sci USA *102*, 13580-13585.

Yao, X., Hu, J.-F., Li, T., Yang, Y., Sun, Z., Ulaner, G.A., Vu, T.H., and Hoffman, A.R. (2004). Epigenetic regulation of the taxol resistance-associated gene TRAG-3 in human tumors. Cancer Genet Cytogenet *151*, 1-13.

Yegnasubramanian, S., Haffner, M.C., Zhang, Y., Gurel, B., Cornish, T.C., Wu, Z., Irizarry, R.A., Morgan, J., Hicks, J., DeWeese, T.L.*, et al.* (2008). DNA hypomethylation arises later in prostate cancer progression than CpG island hypermethylation and contributes to metastatic tumor heterogeneity. Cancer Res *68*, 8954-8967.

Yegnasubramanian, S., Kowalski, J., Gonzalgo, M.L., Zahurak, M., Piantadosi, S., Walsh, P.C., Bova, G.S., De Marzo, A.M., Isaacs, W.B., and Nelson, W.G. (2004). Hypermethylation of CpG islands in primary and metastatic human prostate cancer. Cancer Res *64*, 1975-1986.

Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet *13*, 335-340.

Yuspa, S.H. (1983). Molecular and cellular basis for tumor promotion in mouse skin. Int Symp Princess Takamatsu Cancer Res Fund *14*, 315-326.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev *15*, 2343-2360.

Zhao, C., Blum, J., Chen, A., Kwon, H.Y., Jung, S.H., Cook, J.M., Lagoo, A., and Reya, T. (2007). Loss of beta-catenin impairs the renewal of normal and CML stem cells in vivo. Cancer Cell *12*, 528-541.

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.-J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science *322*, 750-756.

Zilberman, D. (2007). The human promoter methylome. NatGenet *39*, 442-443.

Appendix

Table A1: Literature review detailing studies of genome-wide hypomethylation in cancer

Reference	Global / Gene Specific	Genes / Repeats	Method	Normal Tissue	n=	Cancer	n=	Matched	Methylation	Comments
(Kondo et al.)	Gene	HRAS, KRAS	HpaII/MspI/Southern	Adjacent normal tissue	8	Colonic adenocarcinoma (7) and small cell lung carcinoma	8	Yes	Decrease	
(Weber et al.)	Gene	GH1, CRYGA, CGA and HBG1	Southern	Adjacent normal tissue	23	Colon Benign,Carcinoma	10,13	Yes	Decrease	
(Feinberg and Vogelstein)	Global		HPLC	Adjacent normal colon mucosa	12	Colon Adenoma,Carcinoma	8,11	Yes	Decrease	
(Goelz et al.)	Global		MAA	Adjacent parenchyma	117	Primary invasive carcinomas	117	Yes	Decrease	
(Feinberg et al.)	Global		MAA	Adjacent parenchyma	5	Benign phyllodes tumours	5	Yes	Decrease	
(Soares et al.)	Global		MAA	Adjacent parenchyma	11	Fibroadenomas	11	Yes	Decrease	
(Soares et al.)	Global		MAA	Adjacent parenchyma	3	Sclerosing adenosis	3	Yes	Decrease	
(Soares et al.)	Gene	HERV-K	MspI/HpaII/Southern	Adjacent bladder	13	Bladder; urothelial carcinoma	13	Yes	Decrease	
(Soares et al.)	Repeats	DMHA-1	RLGS & Southern	Adjacent normal liver	21	HCC HBV-	21	Yes	Decrease	

(Santourlidis et al.)	Gene	IGF2	MspI/HpaII/Southern	Adjacent normal kidney	11	Wilms tumour	11	Yes	Decrease	
(Nagai et al.)	Global		Me-C antibody IF	Adjacent tissue	13	Colon adenocarcinoma	13	Yes	Decrease	Intensity of stainir decreased by 17-2
(Sullivan et al.)	Global		MAA	Adjacent normal liver	17	HCC	17	Yes	Decrease	Correlated with gr and size of tumou No benign change
(Hernandez-Blazquez et al.)	Repeats	Sat2 & 3	MspI/HpaII/Southern	Adjacent normal tissue	67	HCC	67	Yes	Decrease	Hypometh in 67% HCC / 18% norma
(Lin et al.)	Repeat	Sat2; D16Z3 locus	MspI/HpaII/BstBI/Southern	Adjacent normal breast	39	Breast	39	Yes	Decrease	Correlation with ty and grade
(Saito et al.)	Repeat	Sat2; D1Z1 locus	BstBI/Southern	Adjacent normal breast	33	Breast	33	Yes	Decrease	
(Tsuda et al.)	Gene	BRCA2	COBRA, BS	Adjacent normal tissue	30	Sporadic epithelial ovarian carcinoma	30	Yes	Decrease	Correlated with elevated mRNA
(Tsuda et al.)	Global		MAA	Adjacent myometria	23	Leiomyoma	23	Yes	Decrease	~2.2 fold increase methyl acceptance leiomyoma.
(Chan et al.)	Global		IHC	Adjacent myometria	23	Leiomyoma	23	Yes	Decrease	
(Li et al.)	Gene	SFN, CLDN4, CLN2, TFF2, S100A4, MSLN, PSCA	MSP	Adjacent normal tissue, microdissected	8	Primary pancreatic carcinoma, microdissected	8	Yes	Decrease	Frequent hypomet all 7 genes (75-10 of cancers). Cells microdissected.
(Li et al.)	Gene	Synuclein gamma	BS	Adjacent normal tissue	5	Breast cancer	5	Yes	Decrease	
(Sato et al.)	Global		MAA	Adjacent normal mucosa	3	Large adenoma	32	Some	Decrease	No change betwee benign or maligna
(Gupta et al.)	Global		MAA	adjacent normal mucosa	22	Colorectal cancer	24	Yes	Decrease	

(Bariol et al.)	Repeat	LINE1	COBRA	Adjacent normal tissue	8	Bladder, head and neck, liver, lung, kidney, prostate, breast, eosophagus, thyroid, lymph node, stomach.	8	Yes	Decrease	
(Bariol et al.)	Gene	Maspin	MSP	Adjacent normal thyroid	17	Poorly differentiated thyroid cancer	17	Yes	Decrease	
(Chalitchagorn et al.)	Gene	Maspin	MSP	Adjacent normal thyroid	13	Undifferentiated thyroid cancer	13	Yes	Decrease	
(Ogasawara et al.)	Gene	TRAG-3	BS, MSP	Adjacent normal tissue	6	1 each of ovary, liver, lung, colon, breast, testis	6	Yes	Decrease	Correlated with expression of TRAG3.
(Ogasawara et al.)	Repeats / Global	L1 promoter	MspI/HpaII/Southern; MAA	adjacent normal colon mucosa	8	Colorectal cancer	8	Yes	Decrease	CRC patient norm bowel displayed greater heterogene than from unaffec individuals.
(Yao et al.)	Gene	Synuclein gamma	MSP, BS	Matched nonneoplastic adjacent tissue	160	Various, 20 each of liver, esophagus, prostate, gastric, colon, cervical, lung and breast	160	Yes	Decrease	49/160 of the adjacent tissues showed demethylation of SNCG.
(Suter et al.)	Global		Me-C antibody IHC	Benign sites within tumour	30	Prostate tumour	30	Some	Decrease	
(Liu et al.)	Repeat	SatR-1	BS & LM-PCR , COBRA	Adjacent normal tissue	11	Primary breast tumour	11	Yes	Decrease	
(Brothman et al.)	Repeat	LINE1	COBRA	Adjacent normal ovary	59	Epithelial ovarian cancer	59	Yes	Decrease	Correlated with ty grade and poorer overall survival.

(Costa et al.)	Gene	GLIPR1/RTVP-1	COBRA	Adjacent normal kidney	17	Wilms tumour	24	Yes	Decrease	
(Pattamadilok et al.)	Repeat	LINE1	MSP	Adjacent normal colon mucosa	178	Colorectal cancer	205	Yes	Decrease	No association between LINE1 ar CIMP phenotype
(Chilukamarri et al.)	Repeat	LINE1 & Alu	BS	Adjacent normal tissue	35	Neuroendocrine tumour	35	Yes	Decrease	Alu showed no change.LINE1 correlated with ile vs non-ileal carcin tumors, with lymp node metastasis, chromosome 18 lo and RASSF1A ge methylation.
(Iacopetta et al.)	Repeat	LINE, SINE	MBD2 based microarray, BS	Normal matched lung tissue	4	Lung SCC (two stage I, one stage II, one stage III)	4	Yes	Decrease	Non-conserved hypomethylation between tumours, suggesting the process is random
(Choi et al.)	Repeats	LINE1	BS	Adjacent normal tissue	26	Head and Neck SCC	26	Yes	Decrease	Association with HPV status
(Rauch et al.)	Gene	HGH, HBG1 and HBA1	HpaII/HhaI/Southern	Adjacent normal tissue	6	Lung & Colon	6	Yes	Decrease / No change	1/4 colon showed change
(Richards et al.)	Repeats	LINE1	MspI/HpaII/Southern & LM-PCR	Adjacent bladder	73	Bladder	73	Yes	Decrease / No change	4/73 showed no change
(Feinberg and Vogelstein)	Repeats	DMHA-1	RLGS & Southern	Adjacent normal liver	8	HCC HBV+	8	Yes	Decrease / No change	2/8 HCC (HBV+), 17/21 (HBV-) showed no change
(Florl et al.)	Repeats	LINE1	MspI/HpaII/Southern	Adjacent normal tissue	9	HCC	9	Yes	Decrease / No change	1/9 showed no change

(Nagai et al.)	Gene	KvDMR1	BamHI/HpaII/Southern	Adjacent normal liver	20	HCC	20	Yes	Decrease / No change	Loss of methylatic in 10/20 randomly selected HCCs.
(Takai et al.)	Gene	WT1	BstuI/Bsh1236I/Southern, BS	Adjacent normal kidney	9	Wilms tumour	9	Yes	Decrease / No change	2/9 showed no change
(Schwienbacher et al.)	Repeats	Sat2	BstBI/Southern	Adjacent normal tissue	36	HCC	36	Yes	Decrease / No change	11/36 showed no change
(Malik et al.)	Gene	MAGE-A1 -A3 and -B2	MSP	Adjacent normal lung	20	Non-small cell lung cancer	20	Yes	Decrease / No change	normal tissue: hypometh in 7/20; 10/20 and 11/20; NSCLC hypometh 15/20, 16/20, 16/2
(Wong et al.)	Gene	CTCF binding site in H19 promoter	BS	Adjacent normal bladder	6	Bladder	6	Yes	Decrease / No change	2/6 cases showed hypomethylation
(Jang et al.)	Global		HPCE	Corresponding normal tissue	57	Carcinoma; lung (n=27), colorectal (n=18), breast (n=12)	57	Yes	Decrease / No change	30% showed no change or higher 5 methylcytosine. MTHFR SNP associated with lo levels of 5-meC in normal tissue.
(Takai et al.)	Gene	LINE1; MAGE genes	HpaII/MspI/Southern; BS	Adjacent normal tissue	2	Gastric cancer	2	Yes	Decrease / No change	1/2 showed no change
(Paz et al.)	Gene	MAGE-A1, -A2, -A3, -B2, -C1, -C2	MSP	Adjacent normal tissue	28	Gastric cancer	93	Some	Decrease / No change	31/93 showed no change
(Kaneda et al.)	Global		HPLC	Adjacent normal tissue	9	Gastric cancer	9	Some	Decrease / No change	5/9 less than lowe level of normal samples

(Kaneda et al.)	Repeats	LINE1, Alu Yb8, Sat-a, NBL-2, D4Z4	BS	Adjacent normal tissue	38	bladder cancer	38	Yes	Decrease / No change	~40% tumors sho no change
(Kaneda et al.)	Repeats	LINE1	MspI/HpaII/Southern & LM-PCR	Adjacent kidney	34	Kidney	34	Yes	No change	
(Choi et al.)	Repeat	Sat2; D16Z3 and D1Z1 loci	MspI/HpaII/BstBI/Southern	Adjacent normal breast	4	Benign breast tumours	4	Yes	No change	
(Florl et al.)	Gene	CAGE	MSP	Normal prostate	14	Prostate	23	Yes	No change	
(Tsuda et al.)	Gene	CAGE	MSP	Normal colon	14	Colorectal	16	Yes	No change	
(Cho et al.)	Gene	Maspin	MSP	Adjacent normal thyroid	6	Follicular thyroid adenoma	6	Yes	No change	Inverse correlatio between maspin promoter methylat and expression.
(Cho et al.)	Gene	Maspin	MSP	Adjacent normal thyroid	56	Well differentiated thyroid cancer	56	Yes	No change	
(Ogasawara et al.)	Repeats	LINE1	BS, LUMA	Adjacent normal tissue	21	follicular thyroid cancer	21	Yes	No change	

Table A2: Probe set IDs of epigenetic regulator genes utilized in GEM analysis

	Probe Set ID	Gene Title	Gene Symbol
	DNA methyltransferases		
DNMT	201697_s_at	DNA (cytosine-5-)-methyltransferase 1	DNMT1
DNMT	206308_at	DNA (cytosine-5-)-methyltransferase 2	DNMT2
DNMT	232121_at	DNA (cytosine-5-)-methyltransferase 2	DNMT2
DNMT	243504_at	DNA (cytosine-5-)-methyltransferase 2	DNMT2
DNMT	218457_s_at	DNA (cytosine-5-)-methyltransferase 3 alpha	DNMT3A
DNMT	222640_at	DNA (cytosine-5-)-methyltransferase 3 alpha	DNMT3A
DNMT	244428_at	DNA (cytosine-5-)-methyltransferase 3 alpha	DNMT3A
DNMT	220668_s_at	DNA (cytosine-5-)-methyltransferase 3 beta	DNMT3B
	Histone deacetylases		
HDAC	201209_at	histone deacetylase 1	HDAC1
HDAC	201833_at	histone deacetylase 2	HDAC2
HDAC	242141_at	Histone deacetylase 2	HDAC2
HDAC	216326_s_at	histone deacetylase 3	HDAC3
HDAC	240482_at	Histone deacetylase 3	HDAC3
HDAC	1554322_a_at	histone deacetylase 4	HDAC4
HDAC	204225_at	histone deacetylase 4	HDAC4
HDAC	228813_at	histone deacetylase 4	HDAC4
HDAC	240870_at	Histone deacetylase 4	HDAC4
HDAC	202455_at	histone deacetylase 5	HDAC5
HDAC	229408_at	histone deacetylase 5	HDAC5
HDAC	206846_s_at	histone deacetylase 6	HDAC6
HDAC	211722_s_at	histone deacetylase 6	HDAC6
HDAC	216224_s_at	histone deacetylase 6	HDAC6
HDAC	217937_s_at	histone deacetylase 7A	HDAC7A
HDAC	236326_at	histone deacetylase 7A	HDAC7A
HDAC	223345_at	histone deacetylase 8	HDAC8
HDAC	223908_at	histone deacetylase 8	HDAC8
HDAC	223909_s_at	histone deacetylase 8	HDAC8
HDAC	1552758_at	histone deacetylase 9	HDAC9
HDAC	1552760_at	histone deacetylase 9	HDAC9
HDAC	205659_at	histone deacetylase 9	HDAC9
HDAC	234393_at	histone deacetylase 9	HDAC9
HDAC	242952_at	Histone deacetylase 9	HDAC9
HDAC	226672_s_at	histone deacetylase 10	HDAC10
HDAC	232870_at	histone deacetylase 10	HDAC10
HDAC	219847_at	histone deacetylase 11	HDAC11
HDAC	227679_at	Histone deacetylase 11	HDAC11
HDAC	218878_s_at	sirtuin (silent mating type information regulation 2 homolog) 1 (S. cerevisiae)	SIRT1
HDAC	1558331_at	sirtuin (silent mating type information regulation 2 homolog) 2 (S. cerevisiae)	SIRT2
HDAC	220605_s_at	sirtuin (silent mating type information regulation 2 homolog) 2 (S. cerevisiae)	SIRT2
HDAC	221562_s_at	sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae)	SIRT3
HDAC	221913_at	sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae)	SIRT3
HDAC	49327_at	sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae)	SIRT3
HDAC	220047_at	sirtuin (silent mating type information regulation 2 homolog) 4 (S. cerevisiae)	SIRT4
HDAC	222248_s_at	sirtuin (silent mating type information regulation 2 homolog) 4 (S. cerevisiae)	SIRT4
HDAC	1569938_at	sirtuin (silent mating type information regulation 2 homolog) 5 (S. cerevisiae)	SIRT5
HDAC	219185_at	sirtuin (silent mating type information regulation 2 homolog) 5 (S. cerevisiae)	SIRT5
HDAC	221010_s_at	sirtuin (silent mating type information regulation 2 homolog) 5 (S. cerevisiae) /// sirtuin (silent mating type information regulation 2 homolog) 5 (S. cerevisiae)	SIRT5

	Probe Set ID	Gene Title	Gene Symbol
HDAC	222080_s_at	Sirtuin (silent mating type information regulation 2 homolog) 5 (S. cerevisiae)	SIRT5
HDAC	222081_at	Sirtuin (silent mating type information regulation 2 homolog) 5 (S. cerevisiae)	SIRT5
HDAC	219613_s_at	sirtuin (silent mating type information regulation 2 homolog) 6 (S. cerevisiae)	SIRT6
HDAC	233179_x_at	sirtuin (silent mating type information regulation 2 homolog) 6 (S. cerevisiae)	SIRT6
HDAC	218797_s_at	sirtuin (silent mating type information regulation 2 homolog) 7 (S. cerevisiae)	SIRT7
HDAC	235046_at	Transcribed locus, moderately similar to XP_518244.1 PREDICTED: similar to sirtuin 5 isoform 2; sir2-like 5; silent mating type information regulation 2, S.cerevisiae, homolog 5; sirtuin (silent mating type information regulation 2, S.cerevisiae, homolog) 5; sirtuin silent mating type information regulation 2 h... [Pan troglodytes]	---
HDAC	239259_at	Transcribed locus, moderately similar to XP_518244.1 PREDICTED: similar to sirtuin 5 isoform 2; sir2-like 5; silent mating type information regulation 2, S.cerevisiae, homolog 5; sirtuin (silent mating type information regulation 2, S.cerevisiae, homolog) 5; sirtuin silent mating type information regulation 2 h... [Pan troglodytes]	---
HDAC	1570574_at	Similar to NAD-dependent deacetylase sirtuin 5 (SIR2-like protein 5)	LOC391047
HDAC	1570575_at	Similar to NAD-dependent deacetylase sirtuin 5 (SIR2-like protein 5)	LOC391047
HDAC	224974_at	SDS3 protein	SDS3
HDAC	233841_s_at	SDS3 protein	SDS3
HDAC	219433_at	BCL6 co-repressor	BCOR
HDAC	223566_s_at	BCL6 co-repressor	BCOR
HDAC	223915_at	BCL6 co-repressor	BCOR
HDAC	223916_s_at	BCL6 co-repressor	BCOR
	Histone acetyltransferases		
HAT	202652_at	amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)	APBB1
HAT	213419_at	amyloid beta (A4) precursor protein-binding, family B, member 2 (Fe65-like)	APBB2
HAT	216747_at	amyloid beta (A4) precursor protein-binding, family B, member 2 (Fe65-like)	APBB2
HAT	216750_at	amyloid beta (A4) precursor protein-binding, family B, member 2 (Fe65-like)	APBB2
HAT	40148_at	amyloid beta (A4) precursor protein-binding, family B, member 2 (Fe65-like)	APBB2
HAT	1559295_at	CREB binding protein (Rubinstein-Taybi syndrome)	CREBBP
HAT	202160_at	CREB binding protein (Rubinstein-Taybi syndrome)	CREBBP
HAT	211808_s_at	CREB binding protein (Rubinstein-Taybi syndrome)	CREBBP
HAT	228177_at	CREB binding protein (Rubinstein-Taybi syndrome)	CREBBP
HAT	235858_at	CREB binding protein (Rubinstein-Taybi syndrome)	CREBBP
HAT	209058_at	endothelial differentiation-related factor 1	EDF1
HAT	209059_s_at	endothelial differentiation-related factor 1	EDF1
HAT	202221_s_at	E1A binding protein p300	EP300
HAT	213579_s_at	E1A binding protein p300	EP300
HAT	200878_at	endothelial PAS domain protein 1	EPAS1
HAT	200879_s_at	endothelial PAS domain protein 1	EPAS1
HAT	229904_at	Endothelial PAS domain protein 1	EPAS1
HAT	235963_at	Endothelial PAS domain protein 1	EPAS1
HAT	237843_at	Endothelial PAS domain protein 1	EPAS1
HAT	241055_at	Endothelial PAS domain protein 1	EPAS1
HAT	242868_at	Endothelial PAS domain protein 1	EPAS1
HAT	202182_at	GCN5 general control of amino-acid synthesis 5-like 2 (yeast)	GCN5L2
HAT	219198_at	general transcription factor IIIC, polypeptide 4, 90kDa	GTF3C4
HAT	203138_at	histone acetyltransferase 1	HAT1
HAT	200989_at	hypoxia-inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)	HIF1A
HAT	238869_at	Hypoxia-inducible factor 1, alpha subunit (basic helix-loop-helix	HIF1A

	Probe Set ID	Gene Title	Gene Symbol
		transcription factor)	
HAT	206689_x_at	HIV-1 Tat interacting protein, 60kDa	HTATIP
HAT	209192_x_at	HIV-1 Tat interacting protein, 60kDa	HTATIP
HAT	214258_x_at	HIV-1 Tat interacting protein, 60kDa	HTATIP
HAT	214885_at	MYST histone acetyltransferase 1	MYST1
HAT	221820_s_at	MYST histone acetyltransferase 1	MYST1
HAT	200049_at	MYST histone acetyltransferase 2 /// MYST histone acetyltransferase 2	MYST2
HAT	1562236_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
HAT	211874_s_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
HAT	212452_x_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
HAT	214496_x_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
HAT	243479_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
HAT	1562314_at	Nuclear receptor coactivator 1	NCOA1
HAT	209105_at	nuclear receptor coactivator 1	NCOA1
HAT	209106_at	nuclear receptor coactivator 1	NCOA1
HAT	209107_x_at	nuclear receptor coactivator 1	NCOA1
HAT	210249_s_at	nuclear receptor coactivator 1	NCOA1
HAT	243826_at	Nuclear receptor coactivator 1	NCOA1
HAT	1562439_at	Nuclear receptor coactivator 3	NCOA3
HAT	207700_s_at	nuclear receptor coactivator 3	NCOA3
HAT	209060_x_at	nuclear receptor coactivator 3	NCOA3
HAT	209061_at	nuclear receptor coactivator 3	NCOA3
HAT	209062_x_at	nuclear receptor coactivator 3	NCOA3
HAT	211352_s_at	nuclear receptor coactivator 3	NCOA3
HAT	203845_at	p300/CBP-associated factor	PCAF
HAT	239585_at	P300/CBP-associated factor	PCAF
HAT	205085_at	origin recognition complex, subunit 1-like (yeast)	ORC1L
HAT	1558331_at	sirtuin (silent mating type information regulation 2 homolog) 2 (S. cerevisiae)	SIRT2
HAT	220605_s_at	sirtuin (silent mating type information regulation 2 homolog) 2 (S. cerevisiae)	SIRT2
HAT	1552630_a_at	Snf2-related CBP activator protein	SRCAP
HAT	1569138_a_at	Snf2-related CBP activator protein	SRCAP
HAT	212275_s_at	Snf2-related CBP activator protein	SRCAP
HAT	213667_at	Snf2-related CBP activator protein	SRCAP
HAT	215053_at	Snf2-related CBP activator protein	SRCAP
HAT	38766_at	Snf2-related CBP activator protein	SRCAP
HAT	216711_s_at	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	TAF1
HAT	216955_at	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	TAF1
HAT	227205_at	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	TAF1
HAT	241982_at	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	TAF1
HAT	1553011_at	TAF1-like RNA polymerase II, TATA box binding protein (TBP)-associated factor, 210kDa	TAF1L
HAT	213299_at	zinc finger and BTB domain containing 7A	ZBTB7A
HAT	213303_x_at	zinc finger and BTB domain containing 7A	ZBTB7A
HAT	219186_at	zinc finger and BTB domain containing 7A	ZBTB7A
HAT	222082_at	zinc finger and BTB domain containing 7A	ZBTB7A
HAT	226554_at	zinc finger and BTB domain containing 7A	ZBTB7A
HAT	230709_x_at	Zinc finger and BTB domain containing 7A	ZBTB7A
HAT	207225_at	arylalkylamine N-acetyltransferase	AANAT
HAT	203025_at	ARD1 homolog A, N-acetyltransferase (S. cerevisiae)	ARD1A
HAT	206964_at	putative N-acetyltransferase Camello 2	CML2
HAT	225432_s_at	CSRP2 binding protein	CSRP2BP
HAT	227316_at	CSRP2 binding protein	CSRP2BP
HAT	228543_at	CSRP2 binding protein	CSRP2BP
HAT	228544_s_at	CSRP2 binding protein	CSRP2BP
HAT	233396_s_at	CSRP2 binding protein	CSRP2BP

	Probe Set ID	Gene Title	Gene Symbol
HAT	218661_at	hypothetical protein FLJ14154	FLJ14154
HAT	45526_g_at	hypothetical protein FLJ14154	FLJ14154
HAT	228880_at	Hypothetical protein LOC339983	FLJ37478
HAT	235316_at	hypothetical protein LOC339983	FLJ37478
HAT	202182_at	GCN5 general control of amino-acid synthesis 5-like 2 (yeast)	GCN5L2
HAT	1562089_at	glycine-N-acyltransferase-like 1	GLYATL1
HAT	227695_at	glycine-N-acyltransferase-like 1	GLYATL1
HAT	227794_at	glycine-N-acyltransferase-like 1	GLYATL1
HAT	225853_at	glucosamine-phosphate N-acetyltransferase 1	GNPNAT1
HAT	217745_s_at	Mak3 homolog (S. cerevisiae)	MAK3
HAT	222393_s_at	Mak3 homolog (S. cerevisiae)	MAK3
HAT	231199_at	Mak3 homolog (S. cerevisiae)	MAK3
HAT	239138_at	Mak3 homolog (S. cerevisiae)	MAK3
HAT	244341_at	Mak3 homolog (S. cerevisiae)	MAK3
HAT	210603_at	hypothetical protein MGC10646	MGC10646
HAT	223040_at	N-acetyltransferase 5 (ARD1 homolog, S. cerevisiae)	NAT5
HAT	210289_at	N-acetyltransferase 8 (camello like)	NAT8
HAT	206963_s_at	N-acetyltransferase 8 (camello like) /// putative N-acetyltransferase Camello 2	NAT8 /// CML2
HAT	204382_at	N-acetyltransferase 9	NAT9
HAT	203845_at	p300/CBP-associated factor	PCAF
HAT	239585_at	P300/CBP-associated factor	PCAF
HAT	206518_s_at	regulator of G-protein signalling 9	RGS9
HAT	203455_s_at	spermidine/spermine N1-acetyltransferase	SAT
HAT	210592_s_at	spermidine/spermine N1-acetyltransferase	SAT
HAT	210593_at	spermidine/spermine N1-acetyltransferase	SAT
HAT	213988_s_at	spermidine/spermine N1-acetyltransferase	SAT
HAT	230333_at	Spermidine/spermine N1-acetyltransferase	SAT
HAT	225272_at	spermidine/spermine N1-acetyltransferase 2	SAT2
	Histone methyltransferases		
HMT	212512_s_at	coactivator-associated arginine methyltransferase 1	CARM1
HMT	214865_at	DOT1-like, histone H3 methyltransferase (S. cerevisiae)	DOT1L
HMT	226201_at	DOT1-like, histone H3 methyltransferase (S. cerevisiae)	DOT1L
HMT	231297_at	DOT1-like, histone H3 methyltransferase (S. cerevisiae)	DOT1L
HMT	219339_s_at	euchromatic histone-lysine N-methyltransferase 1	EHMT1
HMT	222873_s_at	euchromatic histone-lysine N-methyltransferase 1	EHMT1
HMT	225461_at	euchromatic histone-lysine N-methyltransferase 1	EHMT1
HMT	202326_at	euchromatic histone-lysine N-methyltransferase 2	EHMT2
HMT	207484_s_at	euchromatic histone-lysine N-methyltransferase 2	EHMT2
HMT	229079_at	Euchromatic histone-lysine N-methyltransferase 2	EHMT2
HMT	243914_at	Euchromatic histone-lysine N-methyltransferase 2	EHMT2
HMT	203358_s_at	enhancer of zeste homolog 2 (Drosophila)	EZH2
HMT	207772_s_at	HMT1 hnRNP methyltransferase-like 4 (S. cerevisiae)	HRMT1L4
HMT	230839_at	HMT1 hnRNP methyltransferase-like 4 (S. cerevisiae)	HRMT1L4
HMT	213202_at	KIAA0339 gene product	KIAA0339
HMT	1559856_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
HMT	1565034_s_at	AF4/FMR2 family, member 3 /// myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL /// AFF3
HMT	1565254_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila) /// elongation factor RNA polymerase II	MLL /// ELL
HMT	1565436_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
HMT	212076_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
HMT	212078_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
HMT	212079_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
HMT	216624_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
HMT	229935_s_at	Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog,	MLL

	Probe Set ID	Gene Title	Gene Symbol
		Drosophila)	
HMT	1557158_s_at	myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	222413_s_at	myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	222414_at	myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	222415_at	myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	232940_s_at	myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	234182_at	Myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	234651_at	Myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	235859_at	Myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	244010_at	Myeloid/lymphoid or mixed-lineage leukemia 3	MLL3
HMT	1555409_a_at	myeloid/lymphoid or mixed-lineage leukemia 3 /// B melanoma antigen family, member 5 /// B melanoma antigen family, member 3 /// B melanoma antigen family, member 2	MLL3 /// BAGE5 /// BAGE3 /// BAGE2
HMT	219084_at	nuclear receptor binding SET domain protein 1	NSD1
HMT	243612_at	Nuclear receptor binding SET domain protein 1	NSD1
HMT	1556899_at	Protein arginine methyltransferase 5	PRMT5
HMT	216069_at	Protein arginine methyltransferase 2	PRMT2
HMT	216072_at	Protein arginine methyltransferase 2	PRMT2
HMT	241019_at	Protein arginine methyltransferase 3	PRMT3
HMT	219408_at	protein arginine N-methyltransferase 7	PRMT7
HMT	224928_at	SET domain-containing protein 7	SET7
HMT	220200_s_at	PR/SET domain containing protein 8	SET8
HMT	225094_at	PR/SET domain containing protein 8	SET8
HMT	225118_at	PR/SET domain containing protein 8	SET8
HMT	228443_s_at	SET domain containing (lysine methyltransferase) 8	SETD8
HMT	203155_at	SET domain, bifurcated 1	SETDB1
HMT	214197_s_at	SET domain, bifurcated 1	SETDB1
HMT	1564970_at	SET domain, bifurcated 2	SETDB2
HMT	1564972_x_at	SET domain, bifurcated 2	SETDB2
HMT	235338_s_at	SET domain, bifurcated 2	SETDB2
HMT	235339_at	SET domain, bifurcated 2	SETDB2
HMT	238731_at	SET domain, bifurcated 2	SETDB2
HMT	1554059_at	SET domain and mariner transposase fusion gene	SETMAR
HMT	1554060_s_at	SET domain and mariner transposase fusion gene	SETMAR
HMT	206554_x_at	SET domain and mariner transposase fusion gene	SETMAR
HMT	218788_s_at	SET and MYND domain containing 3	SMYD3
HMT	218619_s_at	suppressor of variegation 3-9 homolog 1 (Drosophila)	SUV39H1
HMT	1554572_a_at	suppressor of variegation 3-9 homolog 2 (Drosophila)	SUV39H2
HMT	219262_at	suppressor of variegation 3-9 homolog 2 (Drosophila)	SUV39H2
HMT	1565974_at	Suppressor of variegation 4-20 homolog 1 (Drosophila)	SUV420H1
HMT	1565975_at	Suppressor of variegation 4-20 homolog 1 (Drosophila)	SUV420H1
HMT	1570058_at	suppressor of variegation 4-20 homolog 2 (Drosophila)	SUV420H2
HMT	218242_s_at	suppressor of variegation 4-20 homolog 1 (Drosophila)	SUV420H1
HMT	222566_at	suppressor of variegation 4-20 homolog 1 (Drosophila)	SUV420H1
HMT	222759_at	suppressor of variegation 4-20 homolog 1 (Drosophila)	SUV420H1
HMT	224431_s_at	suppressor of variegation 4-20 homolog 2 (Drosophila) /// suppressor of variegation 4-20 homolog 2 (Drosophila)	SUV420H2
HMT	231457_at	Suppressor of variegation 4-20 homolog 2 (Drosophila)	SUV420H2
HMT	242646_at	Suppressor of variegation 4-20 homolog 1 (Drosophila)	SUV420H1
HMT	243105_at	Suppressor of variegation 4-20 homolog 1 (Drosophila)	SUV420H1
	Histone demethylases		
HDMT	211412_at	peptidyl arginine deiminase, type IV	PADI4
HDMT	211413_s_at	peptidyl arginine deiminase, type IV	PADI4
HDMT	220001_at	peptidyl arginine deiminase, type IV	PADI4
HDMT	1555897_at	amine oxidase (flavin containing) domain 2	AOF2
HDMT	212348_s_at	amine oxidase (flavin containing) domain 2	AOF2
HDMT	238652_at	Amine oxidase (flavin containing) domain 2	AOF2
HDMT	212689_s_at	jumonji domain containing 1A	JMJD1A
HDMT	242758_x_at	Jumonji domain containing 1A	JMJD1A
HDMT	201643_x_at	jumonji domain containing 1B	JMJD1B
HDMT	210878_s_at	jumonji domain containing 1B	JMJD1B

	Probe Set ID	Gene Title	Gene Symbol
HDMT	221763_at	jumonji domain containing 1C	JMJD1C
HDMT	224933_s_at	jumonji domain containing 1C	JMJD1C
HDMT	228793_at	jumonji domain containing 1C	JMJD1C
HDMT	230007_at	jumonji domain containing 1C	JMJD1C
HDMT	241659_at	Jumonji domain containing 1C	JMJD1C
HDMT	241661_at	jumonji domain containing 1C	JMJD1C
HDMT	203204_s_at	jumonji domain containing 2A	JMJD2A
HDMT	203205_at	jumonji domain containing 2A	JMJD2A
HDMT	1558447_at	Jumonji domain containing 2B	JMJD2B
HDMT	1558448_a_at	Jumonji domain containing 2B	JMJD2B
HDMT	1559073_at	Jumonji domain containing 2B	JMJD2B
HDMT	212492_s_at	jumonji domain containing 2B	JMJD2B
HDMT	212495_at	jumonji domain containing 2B	JMJD2B
HDMT	212496_s_at	jumonji domain containing 2B	JMJD2B
HDMT	215616_s_at	jumonji domain containing 2B	JMJD2B
HDMT	216023_at	jumonji domain containing 2B	JMJD2B
HDMT	217664_at	Jumonji domain containing 2B	JMJD2B
HDMT	230126_s_at	Jumonji domain containing 2B	JMJD2B
HDMT	235789_at	jumonji domain containing 2B	JMJD2B
HDMT	240780_at	Jumonji domain containing 2B	JMJD2B
HDMT	1556493_a_at	Jumonji domain containing 2C	JMJD2C
HDMT	209984_at	jumonji domain containing 2C	JMJD2C
HDMT	214861_at	jumonji domain containing 2C	JMJD2C
HDMT	236732_at	Jumonji domain containing 2C	JMJD2C
HDMT	239285_at	Jumonji domain containing 2C	JMJD2C
HDMT	244385_at	Jumonji domain containing 2C	JMJD2C
HDMT	220278_at	jumonji domain containing 2D	JMJD2D
HDMT	242788_at	jumonji domain containing 2D	JMJD2D
HDMT	1556066_at	jumonji domain containing 3	JMJD3
HDMT	1556067_a_at	jumonji domain containing 3	JMJD3
HDMT	218560_s_at	jumonji domain containing 4	JMJD4
HDMT	222671_s_at	jumonji domain containing 4	JMJD4
HDMT	237229_at	Jumonji domain containing 5	JMJD5
	Methyl-C Binding Proteins		
MeCBP	1555611_s_at	methyl-CpG binding domain protein 1	MBD1
MeCBP	203353_s_at	methyl-CpG binding domain protein 1	MBD1
MeCBP	208595_s_at	methyl-CpG binding domain protein 1	MBD1
MeCBP	226862_at	methyl-CpG binding domain protein 1	MBD1
MeCBP	241813_at	methyl-CpG binding domain protein 1	MBD1
MeCBP	202484_s_at	methyl-CpG binding domain protein 2	MBD2
MeCBP	202485_s_at	methyl-CpG binding domain protein 2	MBD2
MeCBP	214396_s_at	methyl-CpG binding domain protein 2	MBD2
MeCBP	214397_at	methyl-CpG binding domain protein 2	MBD2
MeCBP	236130_at	Methyl-CpG binding domain protein 2	MBD2
MeCBP	202463_s_at	methyl-CpG binding domain protein 3	MBD3
MeCBP	41160_at	methyl-CpG binding domain protein 3	MBD3
MeCBP	209579_s_at	methyl-CpG binding domain protein 4	MBD4
MeCBP	209580_s_at	methyl-CpG binding domain protein 4	MBD4
MeCBP	214047_s_at	methyl-CpG binding domain protein 4	MBD4
MeCBP	214048_at	methyl-CpG binding domain protein 4	MBD4
MeCBP	202616_s_at	methyl CpG binding protein 2 (Rett syndrome)	MECP2
MeCBP	202617_s_at	methyl CpG binding protein 2 (Rett syndrome)	MECP2
MeCBP	202618_s_at	methyl CpG binding protein 2 (Rett syndrome)	MECP2
MeCBP	241924_at	Methyl CpG binding protein 2 (Rett syndrome)	MECP2
MeCBP	1564351_at	POU domain, class 2, transcription factor 1	POU2F1
MeCBP	206789_s_at	POU domain, class 2, transcription factor 1	POU2F1
MeCBP	229753_at	POU domain, class 2, transcription factor 1	POU2F1
MeCBP	234649_at	POU domain, class 2, transcription factor 1	POU2F1
MeCBP	214631_at	zinc finger and BTB domain containing 33	ZBTB33
MeCBP	226255_at	zinc finger and BTB domain containing 33	ZBTB33

	Probe Set ID	Gene Title	Gene Symbol
	ATP-dependent chromatin remodelling		
ATP DR	207591_s_at	AT rich interactive domain 1A (SWI- like)	ARID1A
ATP DR	210649_s_at	AT rich interactive domain 1A (SWI- like)	ARID1A
ATP DR	212152_x_at	AT rich interactive domain 1A (SWI- like)	ARID1A
ATP DR	218917_s_at	AT rich interactive domain 1A (SWI- like)	ARID1A
ATP DR	234159_at	AT rich interactive domain 1A (SWI- like)	ARID1A
ATP DR	217985_s_at	bromodomain adjacent to zinc finger domain, 1A	BAZ1A
ATP DR	217986_s_at	bromodomain adjacent to zinc finger domain, 1A	BAZ1A
ATP DR	208806_at	chromodomain helicase DNA binding protein 3	CHD3
ATP DR	208807_s_at	chromodomain helicase DNA binding protein 3	CHD3
ATP DR	201182_s_at	chromodomain helicase DNA binding protein 4	CHD4
ATP DR	201183_s_at	chromodomain helicase DNA binding protein 4	CHD4
ATP DR	201184_s_at	chromodomain helicase DNA binding protein 4	CHD4
ATP DR	231764_at	chromatin accessibility complex 1	CHRAC1
ATP DR	218166_s_at	hepatitis B virus x associated protein	HBXAP
ATP DR	222540_s_at	hepatitis B virus x associated protein	HBXAP
ATP DR	222541_at	hepatitis B virus x associated protein	HBXAP
ATP DR	223818_s_at	hepatitis B virus x associated protein	HBXAP
ATP DR	235381_at	Hepatitis B virus x associated protein /// Homo sapiens, clone IMAGE:3457110, mRNA	HBXAP
ATP DR	209113_s_at	high-mobility group 20B	HMG20B
ATP DR	210719_s_at	high-mobility group 20B	HMG20B
ATP DR	213966_at	High-mobility group 20B	HMG20B
ATP DR	1558581_at	Metastasis associated 1	MTA1
ATP DR	202247_s_at	metastasis associated 1	MTA1
ATP DR	211783_s_at	metastasis associated 1 /// metastasis associated 1	MTA1
ATP DR	203444_s_at	metastasis associated 1 family, member 2	MTA2
ATP DR	208828_at	polymerase (DNA directed), epsilon 3 (p17 subunit)	POLE3
ATP DR	210371_s_at	retinoblastoma binding protein 4	RBBP4
ATP DR	217015_at	retinoblastoma binding protein 4	RBBP4
ATP DR	217301_x_at	retinoblastoma binding protein 4	RBBP4
ATP DR	239071_at	Retinoblastoma binding protein 4	RBBP4
ATP DR	1557954_at	Retinoblastoma binding protein 7	RBBP7
ATP DR	201092_at	retinoblastoma binding protein 7	RBBP7
ATP DR	227520_at	Retinoblastoma binding protein 7	RBBP7
ATP DR	206542_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
ATP DR	206543_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
ATP DR	206544_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
ATP DR	212257_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
ATP DR	212258_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
ATP DR	217707_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
ATP DR	228926_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
ATP DR	241756_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
ATP DR	202983_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 3	SMARCA3
ATP DR	1569073_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	208793_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	208794_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	212520_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	213719_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4

	Probe Set ID	Gene Title	Gene Symbol
ATP DR	213720_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	214360_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	214728_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	215714_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	217656_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	243655_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
ATP DR	202303_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	SMARCA5
ATP DR	213251_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	SMARCA5
ATP DR	213859_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	SMARCA5
ATP DR	212167_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	SMARCB1
ATP DR	228898_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	SMARCB1
ATP DR	231324_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	SMARCB1
ATP DR	201072_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	SMARCC1
ATP DR	201073_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	SMARCC1
ATP DR	201074_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	SMARCC1
ATP DR	201075_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	SMARCC1
ATP DR	239238_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	SMARCC1
ATP DR	201320_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 2	SMARCC2
ATP DR	201321_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 2	SMARCC2
ATP DR	203183_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1	SMARCD1
ATP DR	209518_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1	SMARCD1
ATP DR	201827_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 2	SMARCD2
ATP DR	204099_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3	SMARCD3
ATP DR	231144_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3	SMARCD3
ATP DR	211988_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	SMARCE1
ATP DR	211989_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	SMARCE1
ATP DR	214871_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	SMARCE1
ATP DR	229511_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	SMARCE1
	Imprinted genes		
Imprinted	208481_at	ankyrin repeat and SOCS box-containing 4	ASB4
Imprinted	217228_s_at	ankyrin repeat and SOCS box-containing 4	ASB4
Imprinted	217229_at	ankyrin repeat and SOCS box-containing 4	ASB4
Imprinted	237720_at	ankyrin repeat and SOCS box-containing 4	ASB4
Imprinted	237721_s_at	ankyrin repeat and SOCS box-containing 4	ASB4
Imprinted	237800_at	Ankyrin repeat and SOCS box-containing 4	ASB4
Imprinted	207607_at	achaete-scute complex-like 2 (Drosophila)	ASCL2
Imprinted	229215_at	achaete-scute complex-like 2 (Drosophila)	ASCL2

	Probe Set ID	Gene Title	Gene Symbol
Imprinted	1568743_at	ATPase, Class V, type 10A	ATP10A
Imprinted	214255_at	ATPase, Class V, type 10A	ATP10A
Imprinted	214256_at	ATPase, Class V, type 10A	ATP10A
Imprinted	207886_s_at	calcitonin receptor	CALCR
Imprinted	207887_s_at	calcitonin receptor	CALCR
Imprinted	1561726_s_at	CD81 antigen (target of antiproliferative antibody 1)	CD81
Imprinted	200675_at	CD81 antigen (target of antiproliferative antibody 1)	CD81
Imprinted	213182_x_at	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C
Imprinted	213183_s_at	Cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C
Imprinted	213348_at	Cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C
Imprinted	216894_x_at	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C
Imprinted	219533_at	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C
Imprinted	219534_x_at	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C
Imprinted	222298_at	Coatomer protein complex, subunit gamma 2	COPG2
Imprinted	223457_at	coatomer protein complex, subunit gamma 2	COPG2
Imprinted	236131_at	Coatomer protein complex, subunit gamma 2	COPG2
Imprinted	205832_at	carboxypeptidase A4	CPA4
Imprinted	1554235_at	catenin (cadherin-associated protein), alpha 3	CTNNA3
Imprinted	1556877_at	Catenin (cadherin-associated protein), alpha 3	CTNNA3
Imprinted	220815_at	catenin (cadherin-associated protein), alpha 3	CTNNA3
Imprinted	201893_x_at	decorin	DCN
Imprinted	209335_at	decorin	DCN
Imprinted	211813_x_at	decorin	DCN
Imprinted	211896_s_at	decorin	DCN
Imprinted	234104_at	Decorin	DCN
Imprinted	240556_at	Decorin	DCN
Imprinted	242605_at	Decorin	DCN
Imprinted	207154_at	deiodinase, iodothyronine, type III	DIO3
Imprinted	215506_s_at	DIRAS family, GTP-binding RAS-like 3	DIRAS3
Imprinted	209560_s_at	delta-like 1 homolog (Drosophila)	DLK1
Imprinted	213707_s_at	distal-less homeo box 5	DLX5
Imprinted	1561316_at	Gamma-aminobutyric acid (GABA) A receptor, beta 3	GABRB3
Imprinted	1569689_s_at	gamma-aminobutyric acid (GABA) A receptor, beta 3	GABRB3
Imprinted	205850_s_at	gamma-aminobutyric acid (GABA) A receptor, beta 3	GABRB3
Imprinted	229724_at	gamma-aminobutyric acid (GABA) A receptor, beta 3	GABRB3
Imprinted	227690_at	Gamma-aminobutyric acid (GABA) A receptor, beta 3 /// DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	GABRB3 /// DDX1
Imprinted	227830_at	Gamma-aminobutyric acid (GABA) A receptor, beta 3 /// DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	GABRB3 /// DDX1
Imprinted	200780_x_at	GNAS complex locus	GNAS
Imprinted	200981_x_at	GNAS complex locus	GNAS
Imprinted	211858_x_at	GNAS complex locus	GNAS
Imprinted	212273_x_at	GNAS complex locus	GNAS
Imprinted	214157_at	GNAS complex locus	GNAS
Imprinted	214548_x_at	GNAS complex locus	GNAS
Imprinted	217057_s_at	GNAS complex locus	GNAS
Imprinted	217058_at	GNAS complex locus	GNAS
Imprinted	217673_x_at	GNAS complex locus	GNAS
Imprinted	228173_at	GNAS complex locus	GNAS
Imprinted	229274_at	GNAS complex locus	GNAS
Imprinted	235851_s_at	GNAS complex locus	GNAS
Imprinted	239037_at	GNAS complex locus	GNAS
Imprinted	242816_at	GNAS complex locus	GNAS
Imprinted	242975_s_at	GNAS complex locus	GNAS
Imprinted	209409_at	growth factor receptor-bound protein 10	GRB10
Imprinted	209410_s_at	growth factor receptor-bound protein 10	GRB10
Imprinted	210999_s_at	growth factor receptor-bound protein 10	GRB10
Imprinted	215248_at	growth factor receptor-bound protein 10	GRB10
Imprinted	224646_x_at	H19, imprinted maternally expressed untranslated mRNA	H19
Imprinted	224997_x_at	H19, imprinted maternally expressed untranslated mRNA	H19
Imprinted	214834_at	HBII-437 C/D box snoRNA /// HBII-13 snoRNA	HBII-437 ///

	Probe Set ID	Gene Title	Gene Symbol
			HBII-13
Imprinted	207135_at	5-hydroxytryptamine (serotonin) receptor 2A	HTR2A
Imprinted	211616_s_at	5-hydroxytryptamine (serotonin) receptor 2A /// 5-hydroxytryptamine (serotonin) receptor 2A	HTR2A
Imprinted	215513_at	hydatidiform mole associated and imprinted	HYMAI
Imprinted	202410_x_at	insulin-like growth factor 2 (somatomedin A)	IGF2
Imprinted	210881_s_at	insulin-like growth factor 2 (somatomedin A)	IGF2
Imprinted	201392_s_at	insulin-like growth factor 2 receptor	IGF2R
Imprinted	201393_s_at	insulin-like growth factor 2 receptor	IGF2R
Imprinted	218637_at	hypothetical protein IMPACT	IMPACT
Imprinted	222698_s_at	hypothetical protein IMPACT	IMPACT
Imprinted	206598_at	insulin	INS
Imprinted	213447_at	imprinted in Prader-Willi syndrome	IPW
Imprinted	204486_at	potassium voltage-gated channel, KQT-like subfamily, member 1	KCNQ1
Imprinted	204487_s_at	potassium voltage-gated channel, KQT-like subfamily, member 1	KCNQ1
Imprinted	211217_s_at	potassium voltage-gated channel, KQT-like subfamily, member 1	KCNQ1
Imprinted	220629_at	KCNQ1 downstream neighbor	KCNQ1DN
Imprinted	1557457_at	Potassium voltage-gated channel, KQT-like subfamily, member 1	KCNQ1OT1
Imprinted	1570123_at	Potassium voltage-gated channel, KQT-like subfamily, member 1	KCNQ1OT1
Imprinted	234590_x_at	Potassium voltage-gated channel, KQT-like subfamily, member 1	KCNQ1OT1
Imprinted	237249_at	Potassium voltage-gated channel, KQT-like subfamily, member 1	KCNQ1OT1
Imprinted	243428_at	KCNQ1 overlapping transcript 1	KCNQ1OT1
Imprinted	243435_at	KCNQ1 overlapping transcript 1	KCNQ1OT1
Imprinted	244727_at	KCNQ1 overlapping transcript 1	KCNQ1OT1
Imprinted	206822_s_at	l(3)mbt-like (Drosophila)	L3MBTL
Imprinted	206823_at	l(3)mbt-like (Drosophila)	L3MBTL
Imprinted	210306_at	l(3)mbt-like (Drosophila)	L3MBTL
Imprinted	213837_at	l(3)mbt-like (Drosophila)	L3MBTL
Imprinted	216077_s_at	l(3)mbt-like (Drosophila)	L3MBTL
Imprinted	242637_at	L(3)mbt-like (Drosophila)	L3MBTL
Imprinted	219894_at	MAGE-like 2	MAGEL2
Imprinted	1558144_at	maternally expressed 3	MEG3
Imprinted	210794_s_at	maternally expressed 3	MEG3
Imprinted	212732_at	maternally expressed 3	MEG3
Imprinted	222328_x_at	Maternally expressed 3	MEG3
Imprinted	226210_s_at	maternally expressed 3	MEG3
Imprinted	226211_at	maternally expressed 3	MEG3
Imprinted	227390_at	maternally expressed 3	MEG3
Imprinted	229557_at	maternally expressed 3	MEG3
Imprinted	231529_at	Maternally expressed 3	MEG3
Imprinted	235077_at	maternally expressed 3	MEG3
Imprinted	242246_x_at	Maternally expressed 3	MEG3
Imprinted	231467_at	Maternally expressed (in Callipyge) 8	MEG8
Imprinted	240083_at	Maternally expressed (in Callipyge) 8	MEG8
Imprinted	241260_at	Maternally expressed (in Callipyge) 8	MEG8
Imprinted	202016_at	mesoderm specific transcript homolog (mouse)	MEST
Imprinted	206585_at	makorin, ring finger protein, 3	MKRN3
Imprinted	209550_at	necdin homolog (mouse)	NDN
Imprinted	204239_s_at	neuronatin	NNAT
Imprinted	223464_at	oxysterol binding protein-like 5	OSBPL5
Imprinted	233734_s_at	oxysterol binding protein-like 5	OSBPL5
Imprinted	241467_at	Oxysterol binding protein-like 5	OSBPL5
Imprinted	212092_at	paternally expressed 10	PEG10
Imprinted	212094_at	paternally expressed 10	PEG10
Imprinted	209242_at	paternally expressed 3	PEG3
Imprinted	209243_s_at	paternally expressed 3	PEG3
Imprinted	230068_s_at	Paternally expressed 3	PEG3
Imprinted	209802_at	pleckstrin homology-like domain, family A, member 2	PHLDA2
Imprinted	209803_s_at	pleckstrin homology-like domain, family A, member 2	PHLDA2
Imprinted	229494_s_at	Pleckstrin homology-like domain, family A, member 2	PHLDA2
Imprinted	1559282_at	Pleiomorphic adenoma gene-like 1	PLAGL1

	Probe Set ID	Gene Title	Gene Symbol
Imprinted	207002_s_at	pleiomorphic adenoma gene-like 1	PLAGL1
Imprinted	207943_x_at	pleiomorphic adenoma gene-like 1	PLAGL1
Imprinted	209318_x_at	pleiomorphic adenoma gene-like 1	PLAGL1
Imprinted	244836_at	Pleiomorphic adenoma gene-like 1	PLAGL1
Imprinted	206344_at	paraoxonase 1	PON1
Imprinted	206345_s_at	paraoxonase 1	PON1
Imprinted	221088_s_at	protein phosphatase 1, regulatory (inhibitor) subunit 9A	PPP1R9A
Imprinted	228494_at	Protein phosphatase 1, regulatory (inhibitor) subunit 9A	PPP1R9A
Imprinted	231966_at	protein phosphatase 1, regulatory (inhibitor) subunit 9A	PPP1R9A
Imprinted	233985_x_at	protein phosphatase 1, regulatory (inhibitor) subunit 9A	PPP1R9A
Imprinted	232629_at	prokineticin 2	PROK2
Imprinted	232976_at	Prader-Willi syndrome chromosome region 1	PWCR1
Imprinted	1554992_at	Ras protein-specific guanine nucleotide-releasing factor 1	RASGRF1
Imprinted	210550_s_at	Ras protein-specific guanine nucleotide-releasing factor 1	RASGRF1
Imprinted	215688_at	Ras protein-specific guanine nucleotide-releasing factor 1	RASGRF1
Imprinted	202026_at	succinate dehydrogenase complex, subunit D, integral membrane protein	SDHD
Imprinted	215652_at	succinate dehydrogenase complex, subunit D, integral membrane protein	SDHD
Imprinted	204688_at	sarcoglycan, epsilon	SGCE
Imprinted	204981_at	solute carrier family 22 (organic cation transporter), member 18	SLC22A18
Imprinted	207429_at	solute carrier family 22 (organic cation transporter), member 2	SLC22A2
Imprinted	1570482_at	Solute carrier family 22 (extraneuronal monoamine transporter), member 3	SLC22A3
Imprinted	205421_at	solute carrier family 22 (extraneuronal monoamine transporter), member 3	SLC22A3
Imprinted	242578_x_at	Solute carrier family 22 (extraneuronal monoamine transporter), member 3	SLC22A3
Imprinted	1555345_at	solute carrier family 38, member 4	SLC38A4
Imprinted	220786_s_at	solute carrier family 38, member 4	SLC38A4
Imprinted	1559342_a_at	Small nuclear ribonucleoprotein polypeptide N	SNRPN
Imprinted	1559343_at	Small nuclear ribonucleoprotein polypeptide N	SNRPN
Imprinted	1559545_at	Small nuclear ribonucleoprotein polypeptide N /// Clone RT-24 SNURF-SNRPN mRNA, partial sequence; alternatively spliced	SNRPN
Imprinted	1559546_s_at	Small nuclear ribonucleoprotein polypeptide N /// Clone RT-24 SNURF-SNRPN mRNA, partial sequence; alternatively spliced	SNRPN
Imprinted	1560741_at	Small nuclear ribonucleoprotein polypeptide N	SNRPN
Imprinted	216850_at	small nuclear ribonucleoprotein polypeptide N	SNRPN
Imprinted	221974_at	Small nuclear ribonucleoprotein polypeptide N	SNRPN
Imprinted	226587_at	Small nuclear ribonucleoprotein polypeptide N /// CDNA FLJ33569 fis, clone BRAMY2010317	SNRPN
Imprinted	226591_at	Small nuclear ribonucleoprotein polypeptide N	SNRPN
Imprinted	228370_at	Small nuclear ribonucleoprotein polypeptide N	SNRPN
Imprinted	201522_x_at	small nuclear ribonucleoprotein polypeptide N /// SNRPN upstream reading frame	SNRPN /// SNURF
Imprinted	206042_x_at	small nuclear ribonucleoprotein polypeptide N /// SNRPN upstream reading frame	SNRPN /// SNURF
Imprinted	1552860_at	transcription elongation factor B polypeptide 3C (elongin A3)	TCEB3C
Imprinted	1554379_a_at	tumor protein p73	TP73
Imprinted	220804_s_at	tumor protein p73	TP73
Imprinted	232546_at	Tumor protein p73	TP73
Imprinted	223935_at	transient receptor potential cation channel, subfamily M, member 5	TRPM5
Imprinted	218612_s_at	tumor suppressing subtransferable candidate 4	TSSC4
Imprinted	206512_at	U2(RNU2) small nuclear RNA auxillary factor 1-like 1	U2AF1L1
Imprinted	211285_s_at	ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	UBE3A
Imprinted	211575_s_at	ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	UBE3A
Imprinted	212278_x_at	ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	UBE3A
Imprinted	213128_s_at	ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	UBE3A
Imprinted	213291_s_at	ubiquitin protein ligase E3A (human papilloma virus E6-associated	UBE3A

	Probe Set ID	Gene Title	Gene Symbol
		protein, Angelman syndrome)	
Imprinted	214980_at	Ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	UBE3A
Imprinted	234163_at	Ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	UBE3A
Imprinted	234166_at	Ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	UBE3A
Imprinted	240204_at	Ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	UBE3A
Imprinted	220895_at	ubiquitin specific peptidase 29	USP29
Imprinted	206067_s_at	Wilms tumor 1	WT1
Imprinted	216953_s_at	Wilms tumor 1	WT1
Imprinted	214218_s_at	X (inactive)-specific transcript	XIST
Imprinted	221728_x_at	X (inactive)-specific transcript	XIST
Imprinted	224588_at	X (inactive)-specific transcript	XIST
Imprinted	224589_at	X (inactive)-specific transcript	XIST
Imprinted	224590_at	X (inactive)-specific transcript	XIST
Imprinted	227671_at	X (inactive)-specific transcript	XIST
Imprinted	231592_at	X (inactive)-specific transcript	XIST
Imprinted	235446_at	X (inactive)-specific transcript	XIST
Imprinted	243712_at	X (inactive)-specific transcript	XIST
Imprinted	220653_at	zinc finger, imprinted 2	ZIM2
Imprinted	1553022_at	zinc finger, imprinted 3	ZIM3
Imprinted	1555510_at	zinc finger protein 215	ZNF215
Imprinted	220214_at	zinc finger protein 215	ZNF215
Imprinted	1558698_at	zinc finger protein 264	ZNF264
Imprinted	205917_at	zinc finger protein 264	ZNF264
Imprinted	230063_at	Zinc finger protein 264	ZNF264
	miRNA processing genes		
miRNA	223653_x_at	bruno-like 4, RNA binding protein (Drosophila)	BRUNOL4
miRNA	223654_s_at	bruno-like 4, RNA binding protein (Drosophila)	BRUNOL4
miRNA	231220_at	bruno-like 4, RNA binding protein (Drosophila)	BRUNOL4
miRNA	232719_at	bruno-like 4, RNA binding protein (Drosophila)	BRUNOL4
miRNA	238966_at	Bruno-like 4, RNA binding protein (Drosophila)	BRUNOL4
miRNA	243665_s_at	Bruno-like 4, RNA binding protein (Drosophila)	BRUNOL4
miRNA	243666_at	Bruno-like 4, RNA binding protein (Drosophila)	BRUNOL4
miRNA	1555467_a_at	CUG triplet repeat, RNA binding protein 1	CUGBP1
miRNA	204113_at	CUG triplet repeat, RNA binding protein 1	CUGBP1
miRNA	209489_at	CUG triplet repeat, RNA binding protein 1	CUGBP1
miRNA	221742_at	CUG triplet repeat, RNA binding protein 1	CUGBP1
miRNA	221743_at	CUG triplet repeat, RNA binding protein 1	CUGBP1
miRNA	235297_at	CUG triplet repeat, RNA binding protein 1	CUGBP1
miRNA	235865_at	CUG triplet repeat, RNA binding protein 1	CUGBP1
miRNA	218650_at	DiGeorge syndrome critical region gene 8	DGCR8
miRNA	219811_at	DiGeorge syndrome critical region gene 8	DGCR8
miRNA	64474_g_at	DiGeorge syndrome critical region gene 8	DGCR8
miRNA	91617_at	DiGeorge syndrome critical region gene 8	DGCR8
miRNA	1557063_at	Dicer1, Dcr-1 homolog (Drosophila)	DICER1
miRNA	206061_s_at	Dicer1, Dcr-1 homolog (Drosophila)	DICER1
miRNA	212888_at	Dicer1, Dcr-1 homolog (Drosophila)	DICER1
miRNA	213229_at	Dicer1, Dcr-1 homolog (Drosophila)	DICER1
miRNA	216260_at	Dicer1, Dcr-1 homolog (Drosophila)	DICER1
miRNA	216280_s_at	Dicer1, Dcr-1 homolog (Drosophila)	DICER1
miRNA	216281_at	Dicer1, Dcr-1 homolog (Drosophila)	DICER1
miRNA	218287_s_at	eukaryotic translation initiation factor 2C, 1	EIF2C1
miRNA	222576_s_at	eukaryotic translation initiation factor 2C, 1	EIF2C1
miRNA	219426_at	eukaryotic translation initiation factor 2C, 3	EIF2C3
miRNA	1569408_at	Eukaryotic translation initiation factor 2C, 4	EIF2C4
miRNA	219190_s_at	eukaryotic translation initiation factor 2C, 4	EIF2C4
miRNA	222842_at	Eukaryotic translation initiation factor 2C, 4	EIF2C4
miRNA	227930_at	Eukaryotic translation initiation factor 2C, 4	EIF2C4

	Probe Set ID	Gene Title	Gene Symbol
miRNA	214868_at	piwi-like 1 (Drosophila)	PIWIL1
miRNA	218269_at	ribonuclease III, nuclear	RNASEN
miRNA	241269_at	Ribonuclease III, nuclear	RNASEN
miRNA	223055_s_at	exportin 5	XPO5
miRNA	223056_s_at	exportin 5	XPO5
miRNA	223057_s_at	exportin 5	XPO5
miRNA	233677_at	Exportin 5	XPO5
	Nucleosome assembly		
Nucleosome	234960_at	---	---
Nucleosome	213497_at	ankyrin repeat and BTB (POZ) domain containing 2	ABTB2
Nucleosome	232624_at	ankyrin repeat and BTB (POZ) domain containing 2	ABTB2
Nucleosome	233628_at	Ankyrin repeat and BTB (POZ) domain containing 2	ABTB2
Nucleosome	233691_at	Ankyrin repeat and BTB (POZ) domain containing 2	ABTB2
Nucleosome	203427_at	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)	ASF1A
Nucleosome	203428_s_at	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)	ASF1A
Nucleosome	213561_at	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)	ASF1A
Nucleosome	1555266_a_at	additional sex combs like 2 (Drosophila)	ASXL2
Nucleosome	218659_at	additional sex combs like 2 (Drosophila)	ASXL2
Nucleosome	201518_at	chromobox homolog 1 (HP1 beta homolog Drosophila)	CBX1
Nucleosome	215989_at	chromobox homolog 2 (Pc class homolog, Drosophila)	CBX2
Nucleosome	224138_at	chromobox homolog 2 (Pc class homolog, Drosophila)	CBX2
Nucleosome	226473_at	chromobox homolog 2 (Pc class homolog, Drosophila)	CBX2
Nucleosome	1555920_at	Chromobox homolog 3 (HP1 gamma homolog, Drosophila)	CBX3
Nucleosome	200037_s_at	chromobox homolog 3 (HP1 gamma homolog, Drosophila) /// chromobox homolog 3 (HP1 gamma homolog, Drosophila)	CBX3
Nucleosome	201091_s_at	chromobox homolog 3 (HP1 gamma homolog, Drosophila)	CBX3
Nucleosome	230998_at	Chromobox homolog 3 (HP1 gamma homolog, Drosophila)	CBX3
Nucleosome	206724_at	chromobox homolog 4 (Pc class homolog, Drosophila)	CBX4
Nucleosome	227558_at	chromobox homolog 4 (Pc class homolog, Drosophila)	CBX4
Nucleosome	209715_at	chromobox homolog 5 (HP1 alpha homolog, Drosophila)	CBX5
Nucleosome	212126_at	Chromobox homolog 5 (HP1 alpha homolog, Drosophila)	CBX5
Nucleosome	226085_at	Chromobox homolog 5 (HP1 alpha homolog, Drosophila)	CBX5
Nucleosome	231862_at	Chromobox homolog 5 (HP1 alpha homolog, Drosophila)	CBX5
Nucleosome	234990_at	Chromobox homolog 5 (HP1 alpha homolog, Drosophila)	CBX5
Nucleosome	242069_at	Chromobox homolog 5 (HP1 alpha homolog, Drosophila)	CBX5
Nucleosome	202047_s_at	chromobox homolog 6	CBX6
Nucleosome	202048_s_at	chromobox homolog 6	CBX6
Nucleosome	229733_s_at	Chromobox homolog 6	CBX6
Nucleosome	212914_at	chromobox homolog 7	CBX7
Nucleosome	219755_at	chromobox homolog 8 (Pc class homolog, Drosophila)	CBX8
Nucleosome	204962_s_at	centromere protein A, 17kDa	CENPA
Nucleosome	210821_x_at	centromere protein A, 17kDa	CENPA
Nucleosome	203975_s_at	chromatin assembly factor 1, subunit A (p150)	CHAF1A
Nucleosome	203976_s_at	chromatin assembly factor 1, subunit A (p150)	CHAF1A
Nucleosome	214426_x_at	chromatin assembly factor 1, subunit A (p150)	CHAF1A
Nucleosome	229808_at	Chromatin assembly factor 1, subunit A (p150)	CHAF1A
Nucleosome	204775_at	chromatin assembly factor 1, subunit B (p60)	CHAF1B
Nucleosome	203274_at	coagulation factor VIII-associated (intronic transcript) 1	F8A1
Nucleosome	1552853_at	hypothetical protein FLJ32784	FLJ32784
Nucleosome	1552854_a_at	hypothetical protein FLJ32784	FLJ32784
Nucleosome	232622_at	Hypothetical protein FLJ32784	FLJ32784
Nucleosome	208886_at	H1 histone family, member 0	H1F0
Nucleosome	1553064_at	H1 histone family, member O, oocyte-specific	H1FOO
Nucleosome	204805_s_at	H1 histone family, member X	H1FX
Nucleosome	231004_s_at	H1 histone family, member X	H1FX
Nucleosome	214412_at	H2A histone family, member B3 /// H2A histone family, member B1	H2AFB3 /// H2AFB1
Nucleosome	220936_s_at	H2A histone family, member J	H2AFJ
Nucleosome	224301_x_at	H2A histone family, member J	H2AFJ
Nucleosome	225245_x_at	H2A histone family, member J	H2AFJ
Nucleosome	228213_at	H2A histone family, member J	H2AFJ

	Probe Set ID	Gene Title	Gene Symbol
Nucleosome	202487_s_at	H2A histone family, member V	H2AFV
Nucleosome	212205_at	H2A histone family, member V	H2AFV
Nucleosome	212206_s_at	H2A histone family, member V	H2AFV
Nucleosome	227085_at	H2A histone family, member V	H2AFV
Nucleosome	205436_s_at	H2A histone family, member X	H2AFX
Nucleosome	212524_x_at	H2A histone family, member X	H2AFX
Nucleosome	212525_s_at	H2A histone family, member X	H2AFX
Nucleosome	213344_s_at	H2A histone family, member X	H2AFX
Nucleosome	207168_s_at	H2A histone family, member Y	H2AFY
Nucleosome	214500_at	H2A histone family, member Y	H2AFY
Nucleosome	214501_s_at	H2A histone family, member Y	H2AFY
Nucleosome	226840_at	H2A histone family, member Y	H2AFY
Nucleosome	229593_at	H2A histone family, member Y	H2AFY
Nucleosome	1563319_at	H2A histone family, member Y2	H2AFY2
Nucleosome	218445_at	H2A histone family, member Y2	H2AFY2
Nucleosome	200853_at	H2A histone family, member Z	H2AFZ
Nucleosome	213911_s_at	H2A histone family, member Z	H2AFZ
Nucleosome	208579_x_at	H2B histone family, member S	H2BFS
Nucleosome	208755_x_at	H3 histone, family 3A	H3F3A
Nucleosome	200080_s_at	H3 histone, family 3A /// H3 histone, family 3A /// H3 histone, family 3A pseudogene /// H3 histone, family 3A pseudogene	H3F3A /// LOC440926
Nucleosome	211940_x_at	H3 histone, family 3A /// H3 histone, family 3A pseudogene	H3F3A /// LOC440926
Nucleosome	213828_x_at	H3 histone, family 3A /// H3 histone, family 3A pseudogene	H3F3A /// LOC440926
Nucleosome	209069_s_at	H3 histone, family 3B (H3.3B)	H3F3B
Nucleosome	211997_x_at	H3 histone, family 3B (H3.3B)	H3F3B
Nucleosome	211998_at	H3 histone, family 3B (H3.3B)	H3F3B
Nucleosome	211999_at	H3 histone, family 3B (H3.3B)	H3F3B
Nucleosome	218166_s_at	hepatitis B virus x associated protein	HBXAP
Nucleosome	222540_s_at	hepatitis B virus x associated protein	HBXAP
Nucleosome	222541_at	hepatitis B virus x associated protein	HBXAP
Nucleosome	223818_s_at	hepatitis B virus x associated protein	HBXAP
Nucleosome	235381_at	Hepatitis B virus x associated protein /// Homo sapiens, clone IMAGE:3457110, mRNA	HBXAP
Nucleosome	231229_at	spermatid-specific linker histone H1-like protein	HILS1
Nucleosome	208484_at	histone 1, H1a	HIST1H1A
Nucleosome	214534_at	histone 1, H1b	HIST1H1B
Nucleosome	209398_at	histone 1, H1c	HIST1H1C
Nucleosome	214537_at	histone 1, H1d	HIST1H1D
Nucleosome	208553_at	histone 1, H1e	HIST1H1E
Nucleosome	1553568_a_at	histone 1, H1t	HIST1H1T
Nucleosome	207982_at	histone 1, H1t	HIST1H1T
Nucleosome	208569_at	histone 1, H2ab	HIST1H2AB
Nucleosome	215071_s_at	histone 1, H2ac	HIST1H2AC
Nucleosome	214469_at	histone 1, H2ae	HIST1H2AE
Nucleosome	207156_at	histone 1, H2ag	HIST1H2AG
Nucleosome	214542_x_at	histone 1, H2ai	HIST1H2AI
Nucleosome	208583_x_at	histone 1, H2aj	HIST1H2AJ
Nucleosome	214644_at	histone 1, H2ak	HIST1H2AK
Nucleosome	214554_at	histone 1, H2al	HIST1H2AL
Nucleosome	214481_at	histone 1, H2am	HIST1H2AM
Nucleosome	241519_at	histone 1, H2ba	HIST1H2BA
Nucleosome	208547_at	histone 1, H2bb	HIST1H2BB
Nucleosome	214455_at	histone 1, H2bc	HIST1H2BC
Nucleosome	236193_at	histone 1, H2bc	HIST1H2BC
Nucleosome	209911_x_at	histone 1, H2bd	HIST1H2BD
Nucleosome	222067_x_at	histone 1, H2bd	HIST1H2BD
Nucleosome	235456_at	Histone 1, H2bd	HIST1H2BD
Nucleosome	235681_at	Histone 1, H2bd	HIST1H2BD
Nucleosome	208527_x_at	histone 1, H2be	HIST1H2BE

	Probe Set ID	Gene Title	Gene Symbol
Nucleosome	208490_x_at	histone 1, H2bf	HIST1H2BF
Nucleosome	210387_at	histone 1, H2bg	HIST1H2BG
Nucleosome	215779_s_at	histone 1, H2bg	HIST1H2BG
Nucleosome	208546_x_at	histone 1, H2bh	HIST1H2BH
Nucleosome	208523_x_at	histone 1, H2bi	HIST1H2BI
Nucleosome	214502_at	histone 1, H2bj	HIST1H2BJ
Nucleosome	209806_at	histone 1, H2bk	HIST1H2BK
Nucleosome	207611_at	histone 1, H2bl	HIST1H2BL
Nucleosome	208515_at	histone 1, H2bm	HIST1H2BM
Nucleosome	207226_at	histone 1, H2bn	HIST1H2BN
Nucleosome	214540_at	histone 1, H2bo	HIST1H2BO
Nucleosome	208575_at	histone 1, H3a	HIST1H3A
Nucleosome	208576_s_at	histone 1, H3b	HIST1H3B
Nucleosome	208577_at	histone 1, H3c	HIST1H3C
Nucleosome	214472_at	histone 1, H3d	HIST1H3D
Nucleosome	214522_x_at	histone 1, H3d	HIST1H3D
Nucleosome	239669_at	Histone 1, H3d	HIST1H3D
Nucleosome	214616_at	histone 1, H3e	HIST1H3E
Nucleosome	208506_at	histone 1, H3f	HIST1H3F
Nucleosome	208496_x_at	histone 1, H3g	HIST1H3G
Nucleosome	206110_at	histone 1, H3h	HIST1H3H
Nucleosome	214509_at	histone 1, H3i	HIST1H3I
Nucleosome	214646_at	Histone 1, H3j	HIST1H3J
Nucleosome	208046_at	histone 1, H4a	HIST1H4A
Nucleosome	214516_at	histone 1, H4b	HIST1H4B
Nucleosome	205967_at	histone 1, H4c	HIST1H4C
Nucleosome	208076_at	histone 1, H4d	HIST1H4D
Nucleosome	206951_at	histone 1, H4e	HIST1H4E
Nucleosome	208026_at	histone 1, H4f	HIST1H4F
Nucleosome	208551_at	histone 1, H4g	HIST1H4G
Nucleosome	208180_s_at	histone 1, H4h	HIST1H4H
Nucleosome	208181_at	histone 1, H4h	HIST1H4H
Nucleosome	232035_at	histone 1, H4h	HIST1H4H
Nucleosome	208580_x_at	histone 1, H4k /// histone 1, H4j	HIST1H4K /// HIST1H4J
Nucleosome	214463_x_at	histone 1, H4k /// histone 1, H4j	HIST1H4K /// HIST1H4J
Nucleosome	214562_at	histone 1, H4l	HIST1H4L
Nucleosome	214290_s_at	histone 2, H2aa	HIST2H2AA
Nucleosome	218279_s_at	histone 2, H2aa	HIST2H2AA
Nucleosome	218280_x_at	histone 2, H2aa	HIST2H2AA
Nucleosome	227943_at	Histone 2, H2aa	HIST2H2AA
Nucleosome	202708_s_at	histone 2, H2be	HIST2H2BE
Nucleosome	207046_at	histone 2, H4	HIST2H4
Nucleosome	230738_at	Histone H4/o	HIST2H4
Nucleosome	230795_at	Histone H4/o	HIST2H4
Nucleosome	238529_at	Histone H4/o	HIST2H4
Nucleosome	221582_at	histone 3, H2a	HIST3H2A
Nucleosome	231681_x_at	Histone 3, H2a	HIST3H2A
Nucleosome	208572_at	histone 3, H3	HIST3H3
Nucleosome	208808_s_at	high-mobility group box 2	HMGB2
Nucleosome	236091_at	high-mobility group box 2	HMGB2
Nucleosome	243368_at	High-mobility group box 2	HMGB2
Nucleosome	1554251_at	heterochromatin protein 1, binding protein 3	HP1BP3
Nucleosome	220633_s_at	heterochromatin protein 1, binding protein 3	HP1BP3
Nucleosome	224591_at	heterochromatin protein 1, binding protein 3	HP1BP3
Nucleosome	224592_x_at	heterochromatin protein 1, binding protein 3	HP1BP3
Nucleosome	1559142_at	MYST histone acetyltransferase (monocytic leukemia) 3	MYST3
Nucleosome	202423_at	MYST histone acetyltransferase (monocytic leukemia) 3	MYST3
Nucleosome	216361_s_at	MYST histone acetyltransferase (monocytic leukemia) 3	MYST3
Nucleosome	221715_at	MYST histone acetyltransferase (monocytic leukemia) 3 /// MYST	MYST3

	Probe Set ID	Gene Title	Gene Symbol
		histone acetyltransferase (monocytic leukemia) 3	
Nucleosome	226547_at	MYST histone acetyltransferase (monocytic leukemia) 3	MYST3
Nucleosome	242480_at	MYST histone acetyltransferase (monocytic leukemia) 3	MYST3
Nucleosome	1562236_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
Nucleosome	211874_s_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
Nucleosome	212452_x_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
Nucleosome	214496_x_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
Nucleosome	243479_at	MYST histone acetyltransferase (monocytic leukemia) 4	MYST4
Nucleosome	1556121_at	nucleosome assembly protein 1-like 1	NAP1L1
Nucleosome	204528_s_at	nucleosome assembly protein 1-like 1	NAP1L1
Nucleosome	208752_x_at	nucleosome assembly protein 1-like 1	NAP1L1
Nucleosome	208753_s_at	nucleosome assembly protein 1-like 1	NAP1L1
Nucleosome	208754_s_at	nucleosome assembly protein 1-like 1	NAP1L1
Nucleosome	212967_x_at	nucleosome assembly protein 1-like 1	NAP1L1
Nucleosome	213864_s_at	nucleosome assembly protein 1-like 1	NAP1L1
Nucleosome	219368_at	nucleosome assembly protein 1-like 2	NAP1L2
Nucleosome	204749_at	nucleosome assembly protein 1-like 3	NAP1L3
Nucleosome	1556567_at	nucleosome assembly protein 1-like 4	NAP1L4
Nucleosome	1560339_s_at	nucleosome assembly protein 1-like 4	NAP1L4
Nucleosome	201414_s_at	nucleosome assembly protein 1-like 4	NAP1L4
Nucleosome	220514_at	Nucleosome assembly protein 1-like 4	NAP1L4
Nucleosome	222319_at	Nucleosome assembly protein 1-like 4	NAP1L4
Nucleosome	229505_at	Nucleosome assembly protein 1-like 4	NAP1L4
Nucleosome	236188_s_at	Nucleosome assembly protein 1-like 4	NAP1L4
Nucleosome	244001_at	nucleosome assembly protein 1-like 4	NAP1L4
Nucleosome	228062_at	nucleosome assembly protein 1-like 5	NAP1L5
Nucleosome	228063_s_at	nucleosome assembly protein 1-like 5	NAP1L5
Nucleosome	204412_s_at	neurofilament, heavy polypeptide 200kDa	NEFH
Nucleosome	33767_at	neurofilament, heavy polypeptide 200kDa	NEFH
Nucleosome	1559889_at	Similar to histone H2B histone family	RP5-998N21.6
Nucleosome	200630_x_at	SET translocation (myeloid leukemia-associated)	SET
Nucleosome	200631_s_at	SET translocation (myeloid leukemia-associated)	SET
Nucleosome	210231_x_at	SET translocation (myeloid leukemia-associated)	SET
Nucleosome	213047_x_at	SET translocation (myeloid leukemia-associated)	SET
Nucleosome	213048_s_at	SET translocation (myeloid leukemia-associated)	SET
Nucleosome	40189_at	SET translocation (myeloid leukemia-associated)	SET
Nucleosome	215780_s_at	SET translocation (myeloid leukemia-associated) /// similar to SET protein (Phosphatase 2A inhibitor I2PP2A) (I-2PP2A) (Template activating factor I) (TAF-I) (HLA-DR associated protein II) (PHAPII) (Inhibitor of granzyme A-activated DNase) (IGAAD)	SET /// LOC389168
Nucleosome	1560171_at	SNF2 histone linker PHD RING helicase	SHPRH
Nucleosome	226366_at	SNF2 histone linker PHD RING helicase	SHPRH
Nucleosome	243964_at	SNF2 histone linker PHD RING helicase	SHPRH
Nucleosome	202303_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	SMARCA5
Nucleosome	213251_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	SMARCA5
Nucleosome	213859_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	SMARCA5
Nucleosome	217160_at	testis specific protein, Y-linked 1	TSPY1
Nucleosome	217162_at	testis specific protein, Y-linked 1	TSPY1
Nucleosome	207918_s_at	testis specific protein, Y-linked 1 /// testis specific protein, Y-linked 2	TSPY1 /// TSPY2
Nucleosome	1560647_at	TSPY-like 1	TSPYL1
Nucleosome	1560648_s_at	TSPY-like 1	TSPYL1
Nucleosome	221493_at	TSPY-like 1	TSPYL1
Nucleosome	218012_at	TSPY-like 2	TSPYL2
Nucleosome	233617_at	TSPY-like 3 (pseudogene)	TSPYL3
Nucleosome	212928_at	TSPY-like 4	TSPYL4
Nucleosome	213122_at	TSPY-like 5	TSPYL5
Nucleosome	231339_at	TSPY-like 6	TSPYL6
	Polycomb group	**Trithorax group genes**	

	Probe Set ID	Gene Title	Gene Symbol
PcG TxG	1565034_s_at	AF4/FMR2 family, member 3 /// myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	AFF3 /// MLL
PcG TxG	1556818_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	1557707_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	1558822_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	1566989_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	1566990_x_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	1566991_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	225181_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	225184_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	233339_s_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	238043_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	239674_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	240116_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	244772_at	AT rich interactive domain 1B (SWI1-like)	ARID1B
PcG TxG	1558170_at	POU domain, class 6, transcription factor 1	ASH1L
PcG TxG	218554_s_at	ash1 (absent, small, or homeotic)-like (Drosophila)	ASH1L
PcG TxG	222667_s_at	ash1 (absent, small, or homeotic)-like (Drosophila)	ASH1L
PcG TxG	226447_at	POU domain, class 6, transcription factor 1	ASH1L
PcG TxG	215989_at	chromobox homolog 2 (Pc class homolog, Drosophila)	CBX2
PcG TxG	224138_at	chromobox homolog 2 (Pc class homolog, Drosophila)	CBX2
PcG TxG	226473_at	chromobox homolog 2 (Pc class homolog, Drosophila)	CBX2
PcG TxG	206724_at	chromobox homolog 4 (Pc class homolog, Drosophila)	CBX4
PcG TxG	227558_at	chromobox homolog 4 (Pc class homolog, Drosophila)	CBX4
PcG TxG	219755_at	chromobox homolog 8 (Pc class homolog, Drosophila)	CBX8
PcG TxG	209572_s_at	embryonic ectoderm development	EED
PcG TxG	210656_at	embryonic ectoderm development	EED
PcG TxG	203249_at	enhancer of zeste homolog 1 (Drosophila)	EZH1
PcG TxG	211310_at	enhancer of zeste homolog 1 (Drosophila)	EZH1
PcG TxG	239197_s_at	enhancer of zeste homolog 1 (Drosophila)	EZH1
PcG TxG	239198_at	enhancer of zeste homolog 1 (Drosophila)	EZH1
PcG TxG	32259_at	enhancer of zeste homolog 1 (Drosophila)	EZH1
PcG TxG	203358_s_at	enhancer of zeste homolog 2 (Drosophila)	EZH2
PcG TxG	1559856_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
PcG TxG	1565436_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
PcG TxG	212076_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
PcG TxG	212078_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
PcG TxG	212079_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
PcG TxG	216624_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
PcG TxG	229935_s_at	Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL
PcG TxG	1565254_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila) /// elongation factor RNA polymerase II	MLL /// ELL
PcG TxG	211790_s_at	myeloid/lymphoid or mixed-lineage leukemia 2	MLL2
PcG TxG	227527_at	myeloid/lymphoid or mixed-lineage leukemia 2	MLL2
PcG TxG	227528_s_at	myeloid/lymphoid or mixed-lineage leukemia 2	MLL2
PcG TxG	231974_at	myeloid/lymphoid or mixed-lineage leukemia 2	MLL2
PcG TxG	203792_x_at	polycomb group ring finger 2	PCGF2
PcG TxG	203793_x_at	polycomb group ring finger 2	PCGF2
PcG TxG	213551_x_at	polycomb group ring finger 2	PCGF2
PcG TxG	214239_x_at	polycomb group ring finger 2	PCGF2
PcG TxG	240752_at	Polycomb group ring finger 2	PCGF2
PcG TxG	218338_at	polyhomeotic-like 1 (Drosophila)	PHC1
PcG TxG	225958_at	polyhomeotic-like 1 (Drosophila)	PHC1
PcG TxG	1557852_at	Polyhomeotic-like 2 (Drosophila)	PHC2
PcG TxG	200919_at	polyhomeotic-like 2 (Drosophila)	PHC2

	Probe Set ID	Gene Title	Gene Symbol
PcG TxG	238131_at	polyhomeotic-like 2 (Drosophila)	PHC2
PcG TxG	1552644_a_at	polyhomeotic like 3 (Drosophila)	PHC3
PcG TxG	1567696_at	Polyhomeotic like 3 (Drosophila)	PHC3
PcG TxG	1567697_at	Polyhomeotic like 3 (Drosophila)	PHC3
PcG TxG	1567698_x_at	Polyhomeotic like 3 (Drosophila)	PHC3
PcG TxG	215521_at	polyhomeotic like 3 (Drosophila)	PHC3
PcG TxG	220328_at	polyhomeotic like 3 (Drosophila)	PHC3
PcG TxG	226508_at	polyhomeotic like 3 (Drosophila)	PHC3
PcG TxG	240599_x_at	Polyhomeotic like 3 (Drosophila)	PHC3
PcG TxG	202928_s_at	PHD finger protein 1	PHF1
PcG TxG	40446_at	PHD finger protein 1	PHF1
PcG TxG	208371_s_at	ring finger protein 1	RING1
PcG TxG	35685_at	ring finger protein 1	RING1
PcG TxG	205215_at	ring finger protein 2	RNF2
PcG TxG	1558968_at	Sex comb on midleg-like 1 (Drosophila)	SCML1
PcG TxG	218793_s_at	sex comb on midleg-like 1 (Drosophila)	SCML1
PcG TxG	222747_s_at	sex comb on midleg-like 1 (Drosophila)	SCML1
PcG TxG	206147_x_at	sex comb on midleg-like 2 (Drosophila)	SCML2
PcG TxG	206542_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
PcG TxG	206543_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
PcG TxG	206544_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
PcG TxG	212257_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
PcG TxG	212258_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
PcG TxG	217707_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
PcG TxG	228926_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
PcG TxG	241756_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
PcG TxG	1569073_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	208793_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	208794_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	212520_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	213719_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	213720_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	214360_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	214728_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	215714_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	217656_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	243655_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	SMARCA4
PcG TxG	212167_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	SMARCB1
PcG TxG	228898_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	SMARCB1
PcG TxG	231324_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	SMARCB1
PcG TxG	201320_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 2	SMARCC2
PcG TxG	201321_s_at	SWI/SNF related, matrix associated, actin dependent regulator of	SMARCC2

	Probe Set ID	Gene Title	Gene Symbol
		chromatin, subfamily c, member 2	
PcG TxG	1566190_at	Suppressor of zeste 12 homolog (Drosophila)	SUZ12
PcG TxG	1566191_at	Suppressor of zeste 12 homolog (Drosophila)	SUZ12
PcG TxG	212287_at	suppressor of zeste 12 homolog (Drosophila)	SUZ12
PcG TxG	213971_s_at	suppressor of zeste 12 homolog (Drosophila)	SUZ12
PcG TxG	200047_s_at	YY1 transcription factor /// YY1 transcription factor	YY1
PcG TxG	201901_s_at	YY1 transcription factor	YY1
PcG TxG	201902_s_at	YY1 transcription factor	YY1
PcG TxG	213494_s_at	YY1 transcription factor	YY1
PcG TxG	206182_at	zinc finger protein 134 (clone pHZ-15)	ZNF134
PcG TxG	227729_at	Zinc finger protein 134 (clone pHZ-15)	ZNF134
PcG TxG	240995_at	Zinc finger protein 134 (clone pHZ-15)	ZNF134
	Cytosine deaminases		
C deaminase	219841_at	activation-induced cytidine deaminase	AICDA
C deaminase	224499_s_at	activation-induced cytidine deaminase /// activation-induced cytidine deaminase	AICDA
C deaminase	207158_at	apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1	APOBEC1
C deaminase	206160_at	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 2	APOBEC2
C deaminase	204205_at	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G	APOBEC3G
C deaminase	215579_at	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G	APOBEC3G
C deaminase	214995_s_at	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G /// apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3F	APOBEC3G /// APOBEC3F
C deaminase	205627_at	cytidine deaminase	CDA

Table A3: Significant changes in gene expression between MSC 0 and MSC 5 (expression data not normalized). Log fold changes in expression and FDR corrected p-value (q-value) are indicated for each gene symbol

Gene Symbol	Log Fold Change	q-value	Gene Symbol	Log Fold Change	q-value
PEG10	3.75	0	H2AFZ	0.87	0
HIST1H2BG	2.40	0	XPO5	0.86	1.00E-06
PEG10	1.98	0	CHAF1A	0.86	1.00E-06
HDAC9	1.93	0	NAT5	0.85	0
ORC1L	1.59	0	GRB10	0.84	0
GADD45A	1.45	0	CBX5	0.84	3.40E-05
CHAF1A	1.42	0	CBX5	0.78	2.80E-05
CSRP2BP	1.40	0	DNMT1	0.77	1.00E-06
EZH2	1.35	1.00E-06	GTF3C4	0.77	0
SUV39H2	1.26	0	HMGB2	0.73	7.00E-06
HIST1H2BG	1.25	0	CENPA	0.73	6.00E-06
POLE3	1.24	0	HDAC2	0.72	3.00E-06
H1F0	1.23	0	GNPNAT1	0.71	2.00E-06
SUV420H1	1.17	3.00E-06	GRB10	0.69	1.00E-06
CHAF1A	1.14	0	SMARCA5	0.67	1.00E-06
MLL / ELL	1.07	6.80E-05	SMARCA5	0.67	1.00E-06
MLL / ELL	1.07	6.80E-05	MAK3	0.67	1.00E-06
EED	1.04	0	H2AFZ	0.67	0
HIST1H4C	1.03	1.00E-06	AFF3 / MLL	0.65	0.000118
PHLDA2	1.03	1.00E-06	MLL / AFF3	0.65	0.000118
CSRP2BP	1.00	1.00E-06	BAZ1A	0.65	0.000497
SUV39H1	0.97	0	SUZ12	0.65	0
NEFH	0.95	1.00E-06	NEFH	0.61	0.002574
CHAF1B	0.94	0	MBD1	0.61	6.20E-05
HAT1	0.93	0	NAP1L1	0.59	0.001352
MAK3	0.90	0	XPO5	0.59	5.00E-06
H1FX	0.89	6.00E-05	SMARCA5	0.58	0.000114
NAP1L2	0.89	0	SMARCA5	0.58	0.000114

Gene Symbol	Log Fold Change	q-value
MYST4	-0.59	1.30E-05
MYST4	-0.59	1.30E-05
CHD3	-0.62	5.00E-06
MLL3	-0.64	0.000497
ZNF264	-0.68	3.60E-05
MLL3	-0.69	6.20E-05
SET7	-0.72	0
HMG20B	-0.73	0
EPAS1	-0.75	0
ZBTB7A	-0.75	1.00E-06
NCOA1	-0.78	4.50E-05
MLL3	-0.78	2.00E-06
NCOA3	-0.80	4.00E-06
IPW	-0.81	1.00E-06
H2AFY2	-0.83	1.00E-06
SLC22A18	-0.91	0
NCOA3	-0.93	0
JMJD2B	-0.94	0
H2AFJ	-0.94	0
CDKN1C	-0.95	0
SMARCA2	-0.97	0
SMARCA2	-0.97	0
APOBEC3G / 3F	-0.99	1.00E-06
APOBEC3G	-0.99	3.40E-05
CHD3	-1.01	0

Gene Symbol	Log Fold Change	q-value
NCOA3	-1.01	1.00E-06
NDN	-1.01	0
H2AFJ	-1.02	0
JMJD2B	-1.05	0
SUV420H1	-1.08	0
SNRPN / SNURF	-1.12	0
SMARCA2	-1.12	0
SMARCA2	-1.12	0
SMARCA2	-1.13	0
SMARCA2	-1.13	0
DLX5	-1.17	0
SUV420H1	-1.27	1.00E-06
HIST2H2BE	-1.30	1.00E-06
HDAC4	-1.31	1.00E-06
SNRPN / SNURF	-1.35	0
SNRPN	-1.43	0
HDAC4	-1.44	0
SNRPN	-1.48	1.00E-06
DCN	-2.63	0
DCN	-2.81	0
DCN	-3.01	0
DCN	-3.06	0
NAP1L3	-3.95	0

Table A4: Significant changes in gene expression between MSC 0 and MSC 5 (expression data normalized to expression of PCNA). Log fold changes in expression and FDR corrected p-value (q-value) are indicated for each gene symbol

Gene Symbol	Log Fold Change	q-value	Gene Symbol	Log Fold Change	q-value
NAP1L3	-4.50	0.0021	NCOA1	-1.08	0.0051
DCN	-3.18	0.0000	ATP10A	-1.08	0.0030
HDAC4	-1.86	0.0010	HDAC5	-1.05	0.0015
HIST2H2BE	-1.85	0.0021	SAT2	-1.05	0.0002
DLX5	-1.73	0.0040	MECP2	-1.04	0.0015
SNRPN	-1.67	0.0002	ZBTB7A	-1.02	0.0022
SUV420H1	-1.63	0.0015	JMJD2C	-1.01	0.0019
NDN	-1.57	0.0009	TSPYL1	-1.01	0.0011
KIAA0907	-1.54	0.0049	SIRT2	-1.01	0.0005
CDKN1C	-1.50	0.0002	H3F3A	-1.00	0.0015
H2AFJ	-1.50	0.0000	MYST4	-1.00	0.0077
JMJD2B	-1.49	0.0001	PHC3	-1.00	0.0001
SLC22A18	-1.46	0.0008	SMARCD3	-0.99	0.0006
H2AFY2	-1.38	0.0012	COPG2	-0.98	0.0008
PAR1	-1.36	0.0002	NCOA3	-0.97	0.0007
NCOA3	-1.35	0.0036	NCOA1	-0.96	0.0089
EPAS1	-1.30	0.0000	HP1-BP74	-0.95	0.0026
HMG20B	-1.28	0.0023	SHPRH	-0.95	0.0080
SET7	-1.27	0.0001	MLL	-0.94	0.0056
ZNF264	-1.23	0.0002	MEG3	-0.93	0.0026
MLL3	-1.19	0.0072	CD81	-0.93	0.0009
CHD3	-1.17	0.0038	ABTB2	-0.92	0.0027
SMARCA2	-1.15	0.0084	HIST2H2AA	-0.92	0.0019
SMARCE1	-1.12	0.0005	H3F3A	-0.92	0.0000
HDAC8	-1.08	0.0063	TSPYL3	-0.91	0.0054

Gene Symbol	Log Fold Change	q-value	Gene Symbol	Log Fold Change	q-value
HMG20B	-0.91	0.0022	CPA4	-0.82	0.0033
EIF2C3	-0.90	0.0005	TAF1	-0.82	0.0022
SIRT4	-0.89	0.0003	JMJD2B	-0.82	0.0012
SETMAR	-0.89	0.0014	SHPRH	-0.81	0.0012
JMJD3	-0.89	0.0067	CDKN1C	-0.81	0.0052
PHC1	-0.88	0.0009	GNAS	-0.81	0.0077
EDF1	-0.88	0.0005	SRCAP	-0.80	0.0008
JMJD2A	-0.87	0.0003	GRB10	0.29	0.0015
IGF2R	-0.87	0.0034	NAT5	0.30	0.0002
CREBBP	-0.86	0.0020	HAT1	0.38	0.0001
EIF2C1	-0.86	0.0044	CHAF1B	0.39	0.0046
H2AFX	-0.86	0.0043	PHLDA2	0.47	0.0034
EPAS1	-0.85	0.0069	EED	0.49	0.0004
ASH1L	-0.84	0.0095	POLE3	0.69	0.0059
ARID1B	-0.84	0.0082	SUV39H2	0.71	0.0031
DCN	-0.84	0.0009	CHAF1A	0.87	0.0010
RING1	-0.84	0.0097	GADD45A	0.89	0.0000
KIFC1	-0.83	0.0000	ORC1L	1.04	0.0004
SETDB1	-0.83	0.0005	HDAC9	1.38	0.0057
ZNF134	-0.82	0.0019	HIST1H2BG	1.85	0.0005
PAR1	-0.82	0.0068	PEG10	3.20	0.0012
HIST1H3H	-0.82	0.0075			

Table A5: Gene symbols of PRC2 target genes displaying SUZ12, EZH2 and H3K27me3 occupancy in human ES cells (from Lee *et al.)*

ABCC8	C20orf103	COLEC12	ECEL1
ABTB2	C21orf63	COMP	EFNA1
ADAMTS15	C2orf32	CORO6	EFNA3
ADAMTS18	C3orf15	CRHR1	EGFL6
ADARB2	CA10	CRLF1	EGR3
ADCY4	CACNA1B	CRTAC1	EGR4
ADCY8	CACNA1D	CRYBA2	ELMOD1
ADCYAP1	CACNA1E	CSMD1	EN1
ADRA1A	CACNA1G	CSMD3	EN2
ADRA2A	CALCA	CTNND2	EOMES
ADRB1	CAMK2N1	CX36	EPAS1
ADRB3	CASZ1	CXCL14	EPB41L4A
ALOX15	CBLN1	CXCL16	EPHA5
ALX3	CBLN4	CYP24A1	EPHB1
ALX4	CBR3	CYP26A1	EPHB3
ANKRD19	CBX8	CYP26B1	ERBB4
ANKRD20A	CD34	CYP26C1	ESAM
ANKRD20B	CD8A	CYP27B1	ESPN
ANKRD27	CDH23	DACH1	ESX1L
AQP5	CDH7	DACH2	F2R
ARHGAP20	CDK5R2	DCAMKL2	FAM19A4
ARL9	CDKN2C	DCC	FAM43B
ASCL1	CDX2	DCHS2	FAM5B
ASCL2	CENTA2	DDAH1	FAM5C
ASTN	CGB7	DGKG	FAM80A
ASTN2	CGB8	DGKI	FAM84A
ATBF1	CGI-38	DHH	FBN2
ATF3	CH25H	DIO3	FBP1
ATOH1	CHODL	DKFZP564O0823	FBXL8
ATOH8	CHRD	DKK1	FBXO3
BAPX1	CHRDL2	DKK2	FEV
BARHL1	CHST8	DLL4	FEZ1
BARHL2	CHX10	DLX1	FGF20
BARX1	CIDEA	DLX2	FGF3
BARX2	CITED1	DLX3	FGF5
BCL2	CKLFSF2	DLX4	FGF9
BHLHB3	CLCN5	DMRT1	FIGLA
BHLHB4	CLEC14A	DMRT2	FLI1
BHLHB5	CLSTN2	DMRT3	FLJ11235
BMP8A	CNNM1	DOK6	FLJ13236
BNC1	CNTFR	DPF3	FLJ20032
BTG2	COL24A1	DPY19L2	FLJ32063
C10orf48	COL25A1	DRD5	FLJ32447
C19orf4	COL27A1	DSC3	FLJ33790
C1orf153	COL2A1	DSCAML1	FLJ34922
C1orf32	COL4A5	DUOX1	FLJ35409
C1orf76	COL4A6	DUOX2	FLJ35740
C1orf92	COL9A2	DUSP4	FLJ36166

FLJ37440	GPR120	HS6ST1P	LHX2
FLJ39553	GPR88	HS6ST3	LHX4
FLJ44815	GRIA2	HSF4	LHX5
FLJ45455	GRID1	HSPA6	LHX6
FLJ45983	GRIK1	HTR1A	LHX8
FLJ46347	GRIK3	HTR2C	LMX1B
FLRT2	GRIN3A	HTR7	LOC124842
FOXA2	GRM7	ICAM5	LOC127003
FOXB1	GSC	IGF2AS	LOC143903
FOXD2	GSCL	IGSF21	LOC148898
FOXD3	GSH1	IL1RAPL2	LOC150221
FOXD4b	GSH2	IL7	LOC153684
FOXD4L1	GUCY1A3	INA	LOC200030
FOXD4L2	GUCY2D	INSM2	LOC340529
FOXD4L3	HAND2	INSRR	LOC388394
FOXE1	HBA1	IPF1	LOC388407
FOXF1	HBA2	IRX3	LOC389289
FOXG1B	HES2	IRX4	LOC400120
FOXJ1	HES7	IRX5	LOC405753
FOXL1	HEY1	ISL1	LOC440804
FOXL2	HHAT	ISL2	LOC441413
FRMD3	HHEX	ITGA4	LOC441425
FUT4	HHIP	ITPKA	LOC441426
FZD10	HLX1	JUN	LOC441430
FZD2	HLXB9	KAZALD1	LOC441459
GABRA2	HMX2	KCNA1	LOC56901
GABRA4	HMX3	KCNA3	LOC92162
GAD2	HOXB1	KCNAB1	LPHN3
GALGT	HOXB13	KCNC2	LPL
GALGT2	HOXB2	KCNC4	LRCH2
GALNTL4	HOXB3	KCND3	LRFN5
GALR2	HOXB6	KCNH1	LRP2
GATA2	HOXB7	KCNH3	LRRTM1
GATA3	HOXB8	KCNK12	LTBP2
GATA4	HOXC11	KCNK13	LTK
GATA6	HOXC12	KCNK2	LYSMD2
GBX2	HOXC4	KCNK4	MAB21L1
GDF6	HOXC5	KCNMA1	MAB21L2
GDF7	HOXC6	KCNQ3	MAFB
GDNF	HOXC8	KCNV1	MAL
GHR	HOXD1	KIAA1036	MAPK4
GHSR	HOXD12	KIAA1199	MAPT
GIMAP5	HOXD13	KIAA1324	MCOLN3
GJB2	HOXD3	KIAA1666	MESP1
GLT25D2	HOXD4	KIRREL3	METRNL
GNA14	HOXD8	KL	MGC11324
GPC5	HOXD9	KLF4	MGC26690
GPM6B	HPCAL4	KY	MGC26718
GPR10	HPSE2	LBX1	MGC35555
GPR101	HRK	LGALS3	MGC39545
GPR12	HS3ST3B1	LGR5	MLLT3

MSC	ONECUT2	PRKCE	SLC10A4
MSX1	OPRD1	PRKG1	SLC1A2
MT1A	OSAP	PROK2	SLC1A4
MT1B	OSR1	PTF1A	SLC24A4
MT1H	OTOP1	PTGDR	SLC26A4
MT1K	OTOP2	PTGER2	SLC27A2
MTM	OTOP3	PTGER3	SLC30A2
MYF6	OTP	PTGER4	SLC30A3
MYO5B	OTX1	PTGFR	SLC30A4
MYOD1	OTX2	PTHLH	SLC32A1
NAGS	OXCT2	PTPRT	SLC35F3
NAV2	PAPPA	PTPRU	SLC6A1
NCAM1	PAX1	PXMP2	SLC6A3
NEF3	PAX2	PYY	SLC6A5
NEFL	PAX3	RAB6C	SLC9A2
NELL1	PAX6	RaLP	SLC9A3
NEUROD1	PAX7	RASGRF1	SLCO2A1
NEUROD2	PAX8	RASSF5	SLCO5A1
NEUROG1	PAX9	RAX	SLIT1
NEUROG2	PCDH17	RBP4	SLIT2
NEUROG3	PCDH8	REPS2	SLITRK1
NFIX	PDE4DIP	RGC32	SLITRK3
NIP	PDGFRA	RGS10	SMP3
NKX2-2	PDZK10	RGS20	SNFT
NKX2-3	PDZK3	RGS9BP	SORCS1
NKX2-8	PENK	RIPK3	SORCS3
NKX3-1	PGM5	RNF127	SOX14
NKX6-1	PGR	RNF128	SOX17
NKX6-2	PHOX2A	ROBO3	SOX7
NLF1	PHOX2B	RPS6KA6	SPAG6
NOL4	PIP5K1B	RRP22	SPOCK3
NPAS1	PIR	RSPONDIN	SPON1
NPNT	PITX1	RTN4RL2	SRD5A2
NPR3	PITX2	RYR3	SSTR1
NPTX1	PITX3	SCD5	SSTR2
NPY1R	PKNOX2	SCN4B	ST8SIA2
NR2F2	PKP1	SCNN1G	STK32B
NR4A3	PLEC1	SCTR	STMN2
NRG1	PLXNA2	SEMA6D	STXBP6
NRG2	PMP22	SFRP1	SUSD4
NS5ATP13TP2	PODN	SFRP5	SV2B
NT5C1A	POLE	SGPP2	SYT12
NTN1	POU3F1	SHH	TAL1
NTNG2	POU3F4	SHOX	TBR1
NTRK1	POU4F1	SHOX2	TBX1
NTRK2	POU4F2	SIDT1	TBX2
NXF	POU4F3	SIM2	TBX21
OCA2	PPM1E	SIX1	TBX3
OLFML2B	PRAC	SIX2	TBX5
OLIG2	PRDM12	SIX3	TCEA3
ONECUT1	PRG-3	SIX6	TCF2

TFAP2E	TRH	VDR	ZADH2
THBD	TRIM36	VSX1	ZBTB16
TIP39	TRIM67	WDR8	ZCCHC16
TITF1	TRIM9	WIT-1	ZFHX1B
TLL1	TRPC5	WNT1	ZFYVE28
TLX1	TSLP	WNT10A	ZIC1
TLX2	TTYH1	WNT10B	ZIC4
TMEFF2	UCN	WNT11	ZMYND15
TMEM27	UCP1	WNT16	ZNF312
TMEM30B	UNC5C	WNT2	ZNF436
TMOD2	UNQ9433	WNT3A	ZNF503
TNFSF7	USH1G	WNT6	ZNFN1A3
TP73	VAX1	WNT7A	
TRADD	VAX2	WT1	